

Protein - protein search, using SW model						
on: July 6, 2004, 13:29:24 ; Search time 58.6787 Seconds (without alignments)						
2176.455 Million cell updates/sec						
Title: US-09-857-518A-6						
Score: 2373						
Sequence: 1 MEXIEVINSKTRIKPSTSS.....MAMLEQDPFLALASPKTLI 452						
Scoring table: BLOSUM62						
Gap0 10.0 , Gapext 0.5						
Number of hits satisfying chosen parameters: 1586107						
Minimum DB seq length: 0						
Maximum DB seq length: 2000000000						
Number of summaries: 45						
Listings: First 45 summaries						
Database : A_Geneseq_29Jan04;*						
1: Geneseq_1990s;*						
2: Geneseq_1990s;*						
3: Geneseq_2000s;*						
4: Geneseq_2001s;*						
5: Geneseq_2002s;*						
6: Geneseq_2003as;*						
7: Geneseq_2003bs;*						
8: Geneseq_2004s;*						
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.						
SUMMARIES						
%						
Query	Match	Length	DB	ID	Description	
NC.	Score	NC.	Score	NC.	NC.	NC.
1	2373	100.0	452	3	AAB36439	Aab6439 Standard
2	2373	100.0	452	3	AYT79656	Aay79656 Strawberry
3	2043	86.1	455	3	AAB36454	Aab36454 Strawberry
4	538.5	22.7	726	3	AAG35696	Aag35696 Arabidopsis
5	538.5	22.7	623	3	AAG35695	Aag35695 Arabidopsis
6	513.5	22.5	725	3	AAG35694	Aag35694 Arabidopsis
7	533.5	22.5	823	3	AAG27611	Aag27611 Arabidopsis
8	533.5	22.5	823	3	AAG27610	Aag27610 Arabidopsis
9	533.5	22.5	954	6	AAG27609	Aag27609 Arabidopsis
10	499	21.0	425	3	AAB36458	Aab36458 Lemon, alc
11	498.5	21.0	443	5	ABB92488	Abb92488 Herbicida
12	498	21.0	436	5	AAB91119	Abb91119 Herbicida
13	498	21.0	436	7	ADB91973	Adb91973 Acyltrans
14	485	20.4	446	7	ADB91985	Adb91985 Acyltrans
15	448	18.9	435	7	ADB91986	Adb91986 Acyltrans
16	418	18.5	474	6	BBB2991	Bbb2991 P. somnif
17	438	18.5	474	6	BBB2991	Bbb2991 P. somnif
18	438	18.5	474	6	BBB2997	Bbb2997 P. somnif
19	408.5	17.2	439	7	AYI17403	Aay17403 Clarkia c
20	404.5	17.0	433	2	AYI17400	Aay17400 Clarkia b
21	404.5	17.0	433	6	AYI29783	Aay29783 C. breweri
22	404.5	17.0	433	6	BBB29783	Bbb29783 C. breweri
23	404.5	17.0	433	7	ADB91977	Adb91977 Acyltrans
24	397.5	16.8	439	2	AYI17402	Aay17402 Clarkia c
25	387.5	16.7	439	3	AYI5147	Aay5147 Wheat str

biosynthetic pathway for aliphatic and/or aromatic ester production in fruit. The nucleotide sequences can be inserted into the genome of a fruit-producing plant to regulate aliphatic and/or aromatic ester formation. Aromatic and/or aliphatic esters in microorganisms, plant cells or plants are produced by inserting thiolase, alcohol acyl transferase, alcohol dehydrogenase, pyruvate decarboxylase, aminotransferase and esterase nucleotide sequences into the genome and feeding the microorganism or plant with alcohol, aldehydes, alpha-keto acids, or amino acids and fatty acids, and acyl-CoA. The nucleotides and their proteins can be used in the processed food industry as food additives to enhance the flavour of syrups, ice-creams, frozen desserts, yogurts and confectionery. They are used, as flavouring agents for oral medications and vitamins; provide flavour and aroma in beverages, including alcohol; enhance or reduce fruit flavour, aroma, fragrance or scent; enhance the flavour or aroma of natural, synthetic or artificial products; for the production of novel combinations of artificial flavour substances; as antibacterial or anti-fungal agents; as fragrance or perfumes in cosmetics, creams, sun-protectant products, hair conditions, lengthening agents and fixatives in perfumes, suspensions aids for aluminium salts in anti-perspirant pharmaceuticals, cleaning products, personal care products and animal care products; as insect disinfectant additives; as degreasing solvents for electronics; as insect pheromones; and as dye carriers, solvents, insect repellents, miticides, scabicides. Plasticizers and deodorants. The present sequence represents the specifically claimed strawberry alcohol acyl transferase

Sequence 452 AA;

Query Match 100.0%; Score 2373; DB 3; Length 452;
Best Local Similarity 100.0%; Pred. No. 3_5e-237; Mismatches 0; Indels 0; Gaps 0;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MERKIEVSINSKTHIKPSSSTSPQPKLTLQDQITPPAYPVFFYPTIDHDLNLPTQIA 60
Db 1 MERKIEVSINSKTHIKPSSSTSPQPKLTLQDQITPPAYPVFFYPTIDHDLNLPTQIA 60

Qy 61 DIRQALSETLTYYPLISGRVKNLYIDDFEEGYPYPLLEARVNCDMTDFLRLKIECLNEFV 120
Db 61 DIRQALSETLTYYPLISGRVKNLYIDDFEEGYPYPLLEARVNCDMTDFLRLKIECLNEFV 120

Qy 121 PPKPSMEAISDERYPLIGVQNVFDSGIAIGTVSVSHKLIDGTADEPLKSGAVFRGCR 180
Db 121 PPKPSMEAISDERYPLIGVQNVFDSGIAIGTVSVSHKLIDGTADEPLKSGAVFRGCR 180

Qy 181 ENITHPSSEAALLPPDLDIPEKVDMEALWAGKKAATRFFVGYKTAISSTQDEAKS 240
Db 181 ENITHPSSEAALLPPDLDIPEKVDMEALWAGKKAATRFFVGYKTAISSTQDEAKS 240

Qy 241 ESYVKPSRVAHTVGFKLHLIAASRAALTSGTTSTRLSIAAQAVNLRTMMETVLDNATG 300
Db 241 ESYVKPSRVAHTVGFKLHLIAASRAALTSGTTSTRLSIAAQAVNLRTMMETVLDNATG 300

Qy 301 NLFWQAQALELSEHTTPPEISDLKLCDLYNLLNSVKQNGDYFETPKCKEGRMCEYID 360
Db 301 NLFWQAQALELSEHTTPPEISDLKLCDLYNLLNSVKQNGDYFETPKCKEGRMCEYID 360

Qy 361 FQTMSSMEPAVDYLFLSSWTFNPNLDGWRGRTWAGKTESASCKFILVPTQCGS 420
Db 361 FQTMSSMEPAVDYLFLSSWTFNPNLDGWRGRTWAGKTESASCKFILVPTQCGS 420

Qy 421 GIEAWVNLEBEKMANLBDQDPHIALASPTKL 452
Db 421 GIEAWVNLEBEKMANLBDQDPHIALASPTKL 452

XX DE Strawberry alcohol acyl transferase SLE27.
XX DE Strawberry; alcohol acyl transferase; fruit; ripening; ester; flavour;
CC KRW aroma; transgenic plant.
XX OS *Frageria x ananassa*.
XX PN EP1003190-A1.
XX PD 07-JUN-2000.
XX PP 02-DEC-1998; 98EP-00234018.
XX PR 02-DEC-1998; 98EP-0024018.
XX (CPRO)-CPRO-DLO CENT PLANTENVERDELINGS REPROD.
PI Verhoeven HA, Van Tunen AJ, Aharoni A, Luecker J, O'connell AP;
DR WPL: 2000-78264/33.
DR N-PSDB, AAA27666
XX PI New polynucleotides encoding enzymes from the biosynthetic pathway for aromatic and/or aliphatic ester production in fruit used to modify plant flavors.
XX PS Claim 7; Page 60-62; 116pp; English.
XX The present sequence is that of strawberry cv. Elsanta alcohol acyl transferase SLE27. Expression of SLE27 increases in ripening fruit, with highest levels in turning and red fruit when volatile ester compounds are at their maximum. The invention relates to DNA sequences (see AAA27666-CC 78) encoding enzymes (see PAV29656-18) involved in the metabolic pathway leading to the formation of aliphatic and/or aromatic esters in ripening fruit. The enzymes have alcohol acyl transferase, alcohol dehydrogenase, pyruvate decarboxylase, thiolase or aminotransferase activity. Expression vectors comprising the DNA sequences may be used to regulate ester production in fruit. Genetically modified plants, plant cells and microorganisms can be used to produce esters. The DNA sequences, polypeptides and antibodies are also used to screen fruit; for volatile ester compounds; for quality such as flavour, fragrance, aroma, scent, texture or shape; to distinguish between cultivars and varieties; and to monitor harvest time, post-harvest quality, shelf-life, timing of pesticide application, and resistance capacity based on volatile ester profiles. (Updated on 12-SEP-2003 to standardise OS field)

Sequence 452 AA;

Query Match 100.0%; Score 2373; DB 3; Length 452;
Best Local Similarity 100.0%; Pred. No. 3_5e-237; Mismatches 0; Indels 0; Gaps 0;
Matches 452; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MERKIEVSINSKTHIKPSSSTSPQPKLTLQDQITPPAYPVFFYPTIDHDLNLPTQIA 60
Db 1 MERKIEVSINSKTHIKPSSSTSPQPKLTLQDQITPPAYPVFFYPTIDHDLNLPTQIA 60

Qy 61 DIRQALSETLTYYPLISGRVKNLYIDDFEEGYPYPLLEARVNCDMTDFLRLKIECLNEFV 120
Db 61 DIRQALSETLTYYPLISGRVKNLYIDDFEEGYPYPLLEARVNCDMTDFLRLKIECLNEFV 120

Qy 121 PPKPSMEAISDERYPLIGVQNVFDSGIAIGTVSVSHKLIDGTADEPLKSGAVFRGCR 180
Db 121 PPKPSMEAISDERYPLIGVQNVFDSGIAIGTVSVSHKLIDGTADEPLKSGAVFRGCR 180

Qy 181 ENITHPSSEAALLPPDLDIPEKVDMEALWAGKKAATRFFVGYKTAISSTQDEAKS 240
Db 181 ENITHPSSEAALLPPDLDIPEKVDMEALWAGKKAATRFFVGYKTAISSTQDEAKS 240

Qy 241 ESYVKPSRVAHTVGFKLHLIAASRAALTSGTTSTRLSIAAQAVNLRTMMETVLDNATG 300
Db 241 ESYVKPSRVAHTVGFKLHLIAASRAALTSGTTSTRLSIAAQAVNLRTMMETVLDNATG 300

Qy 301 NLFWQAQALELSEHTTPPEISDLKLCDLYNLLNSVKQNGDYFETPKCKEGRMCEYID 360
Db 301 NLFWQAQALELSEHTTPPEISDLKLCDLYNLLNSVKQNGDYFETPKCKEGRMCEYID 360

Qy 361 FQTMSSMEPAVDYLFLSSWTFNPNLDGWRGRTWAGKTESASCKFILVPTQCGS 420
Db 361 FQTMSSMEPAVDYLFLSSWTFNPNLDGWRGRTWAGKTESASCKFILVPTQCGS 420

Qy 421 GIEAWVNLEBEKMANLBDQDPHIALASPTKL 452
Db 421 GIEAWVNLEBEKMANLBDQDPHIALASPTKL 452

RESULT 2
AAV79656
ID AAV79656 standard; protein; 452 AA.
XX AC AAV79656;
XX DT 12-SEP-2003 (revised)
DT 29-AUG-2003 (first entry)

301 NLFMFAQILELSTHTEPEISDLKLCDLNLLNGSYQCGNGYFETEKGRGKGRGYRCYELD 360
 301 NLFMFAQILELSTHTEPEISDLKLCDLNLLNGSYQCGNGYFETEKGRGKGRGYRCYELD 360
 Db 361 FORTISSMPEPDIVLFFSWNTNFNPLDGRGRTSMWIGVACKIESASCKEFLIVPTQCGS 420
 Qy 361 FORTISSMPEPDIVLFFSWNTNFNPLDGRGRTSMWIGVACKIESASCKEFLIVPTQCGS 420
 Db 421 GIEAWNLEBEXMAMLECDPFLALASPTKII 452
 Qy 421 GIEAWNLEBEXMAMLECDPFLALASPTKII 452
 Db 421 GIEAWNLEBEXMAMLECDPFLALASPTKII 452

RESULT 3
 ID AAB36454 standard; protein: 455 AA.

XX AAB36454;

AC AAB36454;

XX DT 28-FEB-2001 (first entry)

Strawberry vesca alcohol acyl transferase protein SEQ ID NO:168.
 DE Strawberry vesca alcohol acyl transferase; aliphatic; thiolase;
 XX aromatic ester alcohol acyl transferase; alcohol dehydrogenase;
 KW pyruvate decarboxylase; aminotransferase; alcohol; aldehyde;
 KW alpha-keto acid; amino acid; fatty acid; acyl-CoA; processed food;
 KW food additive; flavour; syrup; ice-cream; frozen dessert; Yoghurt;
 KW confectionery; flavouring; oral medication; vitamin; aroma; beverage;
 KW alcohol; scent; fragrance; perfume; cosmetic; suspension aid;
 KW aluminium salt; anti-perspirant; pharmaceutical; cleaning product;
 KW insect pheromone; dye carrier; solvent; insect repellent; miticide;
 KW scabicide; plasticiser; deodorant.
 XX Fragaria x ananassa.
 XX WO200327379-A1.
 XX PD 08-JUN-2000.
 XX PF 02-DEC-1999; 99WO-NL000737.
 XX PR 02-DEC-1998; 98EP-00204018.
 PR 12-MAR-1999; 99EP-00200739.
 XX PA (CPR0-) CPRO-DLO CENT PLANTENVERDELINGS REPRD.

Pi Aharoni A, Luecker J, Verhoeven H, Van Tunen AJ, O' connell AP;
 XX WPI: 2000-412335/35.
 DR N-PSBB; AAC64783.
 XX A new DNA sequence encoding a polypeptide with alcohol acyl transferase activity for producing and regulating aromatic and/or aliphatic ester formation in microorganisms, plant cells or plants.
 XX PS Example 5; Page 107; 163pp; English.

XX The present invention describes nucleotide sequences with thiolase, alcohol acyl transferase, alcohol dehydrogenase, pyruvate decarboxylase, aminotransferase and esterase activities, which are involved in the biosynthetic pathway for aliphatic and/or aromatic ester production in fruit. The nucleotide sequences can be inserted into the genome of a fruit-producing plant to regulate aliphatic and/or aromatic ester formation. Aromatic and/or aliphatic esters in microorganisms, plant cells or plants are produced by inserting thiolase, alcohol acyl transferase, alcohol dehydrogenase, pyruvate decarboxylase, aminotransferase and esterase nucleotide sequences into the genome and feeding the microorganism or plant with alcohol, aldehydes, alpha-keto acids, or amino acids and fatty acids, and acyl-CoA. The nucleotides and their proteins can be used in the processed food industry as food additives to enhance the flavour of syrups, ice-creams, frozen desserts, yoghurts and confectionery. They are used: as flavouring agents for oral

medications and vitamins; provide flavour and aroma in beverages, including alcohol; enhance or reduce fruit flavour, aroma, fragrance or scent; enhance the flavour or aroma of natural, synthetic or artificial products; for the production of novel combinations of natural, synthetic or artificial flavour substances; as antibacterial or anti-fungal agents; as fragrance or perfumes in cosmetics, creams, sun-protectant products, hair conditioners, lengthening agents and fixatives in perfumes, suspensions aids for aluminium salts in anti-perspirant pharmaceuticals, cleaning products, personal care products and animal care products; as disinfectant additives; as degreasing solvents for electronics; as insect pheromones; and as dye carriers, solvents, insect repellents, miticides, plasticides, plasticisers and decolorants. The present sequence represents strawberry vesca alcohol acyl transferase, from the present invention

XX Sequence 455 AA;
 Query Match 86.1%; Score 2043; DB 3; Length 455;
 Best Local Similarity 87.6%; Pred. No. 7.1e-203;
 Matches 397; Conservative 15; Mismatches 39; Indels 2; Gaps 2;
 CC 1 MEKIEUSINSKHTIKPDTSSPQLQPKMUTIDQLTPAYVEVFFYFPITDHD-FNLPTQL 59
 CC 1 NEKIEVIIISKHTIKPDTSSPQLQPKTLLIDQLTPSYVENVFFYITGPAVFNL-QTL 59
 Qy 60 ADLQRQASBETLTYPPISGRVKNLYIDDF2E2GVPYLEARYNCDMDFLRLKIECNEF 119
 Db 60 ADLQRQASBETLTYPPISGRVKNLYIDDF2E2GVPYLEARYNCDMDFLRLKIECNEF 119
 Qy 120 VP1KPFMEASIDERYLLGQVNWFGSGIAIGVSYSHKLIDGGTACFLKSGWAVPRGC 179
 Db 120 VP1KPFMEASIDERYLLGQVNWFGSGIAIGVSYSHKLIDGGTACFLKSGWAVRGS 179
 Qy 180 RENLTHPSLSEAALEPPRDLPEKYDVKDOMEALWFAKGVKATRFRVFGVKAISS1QDEAK 239
 Db 180 RUKIHNLSQALLPPRDLPEKYDVKDOMEALWFAKGVKATRFRVFGAKAISVQDAK 239
 Qy 240 SESVPKPSRVQHATVGTPLWQHLLIAASPL2LTSGTSTRUSTIAAQAVNLTRMMETVLDNAT 299
 Db 240 SESVPKPSRVQHATVGTPLWQHLLIAASPL2LTSGTSTRUSTIAAQAVNLTRMMETVLDNAT 299
 Qy 300 GNLWFAQAILESEHTPEISDLKLCDLNLLNGSYQCGNGYFETEKGRGKGRGYRCYELD 359
 Db 300 GNLWFAQAILESEHTPEISDLKLCDLNLLNGSYQCGNGYFETEKGRGKGRGYRCYELD 359
 Qy 360 D9RTMSMSEPAFDIYLELESSNTNFNPFLDGFGRTSWIGVAGKIESASCKFILLVPTQCG 419
 Db 360 D9RTMSMSEPAFDIYLELESSNTNFNPFLDGFGRTSWIGVAGKIESASCKFILLVPTQCG 419
 Qy 420 SGTEAWNLEEKKAMIEQDOPHFLALASPTKII 452
 Db 420 TGTEAWNLEEKKAMIEQDOPHFLALASPTKII 452
 XX RESULT 4
 XX AAG39696 standard; protein: 726 AA.
 XX AC AAG39696;
 XX DT 18-OCT-2000 (first entry)
 XX DB Arabidopsis thaliana protein fragment SEQ ID NO: 49156.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridization assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033415-A2.
 XX PD 06-SEP-2000.

25-FEB-2000; 2000EP-00301439.

PF XX 25-FEB-1999; 99US-01211825P.
 PR 01-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 29-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 08-APR-1999; 99US-0128234P.
 PR 16-APR-1999; 99US-0128714P.
 PR 19-APR-1999; 99US-0129845P.
 PR 21-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 22-APR-1999; 99US-0130510P.
 PR 22-APR-1999; 99US-0130891P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 04-MAY-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132488P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0132956P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 19-MAY-1999; 99US-0134768P.
 PR 20-MAY-1999; 99US-0134941P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 28-MAY-1999; 99US-0136392P.
 PR 01-JUN-1999; 99US-0136782P.
 PR 10-JUN-1999; 99US-0137222P.
 PR 14-JUN-1999; 99US-0137528P.
 PR 07-JUN-1999; 99US-0137502P.
 PR 16-JUN-1999; 99US-0137724P.
 PR 18-JUN-1999; 99US-0138094P.
 PR 18-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 16-JUN-1999; 99US-0139119P.
 PR 18-JUN-1999; 99US-0139452P.
 PR 18-JUN-1999; 99US-0139453P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 21-JUN-1999; 99US-0139460P.
 PR 22-JUN-1999; 99US-0139467P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139453P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 28-JUN-1999; 99US-0139617P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-014390P.
 PR 08-JUL-1999; 99US-0142803P.

PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143545P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144088P.
 PR 16-JUL-1999; 99US-0144088P.
 PR 19-JUL-1999; 99US-0144328P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145277P.
 PR 22-JUL-1999; 99US-0145910P.
 PR 22-JUL-1999; 99US-0145910P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0146388P.
 PR 26-JUL-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147039P.
 PR 03-AUG-1999; 99US-0147104P.
 PR 04-AUG-1999; 99US-0147115P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147267P.
 PR 05-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147156P.
 PR 06-AUG-1999; 99US-0147156P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148344P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149422P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 10-SEP-1999; 99US-015166P.
 PR 13-SEP-1999; 99US-0151080P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 25-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152361P.
 PR 10-SEP-1999; 99US-015166P.
 PR 13-SEP-1999; 99US-015375BP.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154032P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155132P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.

18-JUN-1999; 990US-0139458P.
 PR 18-JUN-1999; 990US-0139459P.
 PR 18-JUN-1999; 990US-0139460P.
 PR 18-JUN-1999; 990US-0139461P.
 PR 18-JUN-1999; 990US-0139462P.
 PR 18-JUN-1999; 990US-0139463P.
 PR 18-JUN-1999; 990US-01394750P.
 PR 18-JUN-1999; 990US-0139763P.
 PR 21-JUN-1999; 990US-0139817P.
 PR 22-JUN-1999; 990US-0139899P.
 PR 23-JUN-1999; 990US-0140353P.
 PR 24-JUN-1999; 990US-0140354P.
 PR 24-JUN-1999; 990US-0140695P.
 PR 28-JUN-1999; 990US-0140823P.
 PR 21-JUL-1999; 990US-0140991P.
 PR 30-JUN-1999; 990US-0141287P.
 PR 01-JUL-1999; 990US-0141842P.
 PR 12-JUL-1999; 990US-0142920P.
 PR 13-JUL-1999; 990US-0142154P.
 PR 02-JUL-1999; 990US-0142055P.
 PR 06-JUL-1999; 990US-0142390P.
 PR 08-JUL-1999; 990US-0142803P.
 PR 09-JUL-1999; 990US-0143287P.
 PR 12-JUL-1999; 990US-0142977P.
 PR 14-JUL-1999; 990US-0143542P.
 PR 14-JUL-1999; 990US-0143624P.
 PR 15-JUL-1999; 990US-0144005P.
 PR 15-JUL-1999; 990US-0144085P.
 PR 16-JUL-1999; 990US-0144086P.
 PR 19-JUL-1999; 990US-014415P.
 PR 20-JUL-1999; 990US-0144331P.
 PR 19-JUL-1999; 990US-0144332P.
 PR 20-JUL-1999; 990US-0144332P.
 PR 19-JUL-1999; 990US-0144332P.
 PR 19-JUL-1999; 990US-0144334P.
 PR 21-JUL-1999; 990US-0144335P.
 PR 22-JUL-1999; 990US-0144352P.
 PR 19-JUL-1999; 990US-01444332P.
 PR 20-JUL-1999; 990US-01444884P.
 PR 21-JUL-1999; 990US-0144814P.
 PR 21-JUL-1999; 990US-0145086P.
 PR 21-JUL-1999; 990US-0145088P.
 PR 22-JUL-1999; 990US-0145085P.
 PR 22-JUL-1999; 990US-0145087P.
 PR 22-JUL-1999; 990US-0145089P.
 PR 22-JUL-1999; 990US-0145192P.
 PR 22-JUL-1999; 990US-0145145P.
 PR 23-JUL-1999; 990US-0145218P.
 PR 23-JUL-1999; 990US-0145224P.
 PR 26-JUL-1999; 990US-014566P.
 PR 27-JUL-1999; 990US-0145913P.
 PR 27-JUL-1999; 990US-0145918P.
 PR 27-JUL-1999; 990US-0145919P.
 PR 28-JUL-1999; 990US-0145951P.
 PR 02-AUG-1999; 990US-0146386P.
 PR 02-AUG-1999; 990US-0146388P.
 PR 03-AUG-1999; 990US-0146389P.
 PR 04-AUG-1999; 990US-0147038P.
 PR 04-AUG-1999; 990US-0147204P.
 PR 04-AUG-1999; 990US-0147302P.
 PR 05-AUG-1999; 990US-0147303P.
 PR 06-AUG-1999; 990US-0147260P.
 PR 09-AUG-1999; 990US-0147493P.
 PR 09-AUG-1999; 990US-0147935P.
 PR 10-AUG-1999; 990US-0148171P.
 PR 11-AUG-1999; 990US-0148319P.
 PR 12-AUG-1999; 990US-0148341P.
 PR 13-AUG-1999; 990US-0148565P.
 PR 16-AUG-1999; 990US-0148684P.
 PR 16-AUG-1999; 990US-0149368P.
 PR 18-AUG-1999; 990US-0149426P.
 PR 20-AUG-1999; 990US-0149722P.

PR 20-AUG-1999; 990US-0149723P.
 PR 20-AUG-1999; 990US-0149929P.
 PR 23-AUG-1999; 990US-0149902P.
 PR 23-AUG-1999; 990US-0149930P.
 PR 25-AUG-1999; 990US-0150566P.
 PR 26-AUG-1999; 990US-0150884P.
 PR 27-AUG-1999; 990US-0151055P.
 PR 27-AUG-1999; 990US-0151066P.
 PR 27-AUG-1999; 990US-0151080P.
 PR 30-AUG-1999; 990US-0151303P.
 PR 31-AUG-1999; 990US-0151438P.
 PR 01-SEP-1999; 990US-0151930P.
 PR 07-SEP-1999; 990US-0152363P.
 PR 10-SEP-1999; 990US-0153070P.
 PR 13-SEP-1999; 990US-0153758P.
 PR 15-SEP-1999; 990US-0154018P.
 PR 16-SEP-1999; 990US-0154039P.
 PR 20-SEP-1999; 990US-0155779P.
 PR 22-SEP-1999; 990US-0155139P.
 PR 23-SEP-1999; 990US-0155486P.
 PR 24-SEP-1999; 990US-015659P.
 PR 28-SEP-1999; 990US-0158232P.
 PR 29-SEP-1999; 990US-0158458P.
 PR 04-OCT-1999; 990US-015856P.
 PR 05-OCT-1999; 990US-0157117P.
 PR 06-OCT-1999; 990US-0157753P.
 PR 07-OCT-1999; 990US-0158029P.
 PR 08-OCT-1999; 990US-0158232P.
 PR 12-OCT-1999; 990US-0158499P.
 PR 13-OCT-1999; 990US-0159293P.
 PR 13-OCT-1999; 990US-0159294P.
 PR 13-OCT-1999; 990US-0159295P.
 PR 14-OCT-1999; 990US-0159329P.
 PR 14-OCT-1999; 990US-0159330P.
 PR 14-OCT-1999; 990US-0159331P.
 PR 14-OCT-1999; 990US-0159637P.
 PR 14-OCT-1999; 990US-0159638P.
 PR 18-OCT-1999; 990US-0159584P.
 PR 21-OCT-1999; 990US-0160741P.
 PR 21-OCT-1999; 990US-0160767P.
 PR 21-OCT-1999; 990US-0160789P.
 PR 21-OCT-1999; 990US-0160770P.
 PR 21-OCT-1999; 990US-0160814P.
 PR 21-OCT-1999; 990US-0160815P.
 PR 22-OCT-1999; 990US-0160980P.
 PR 22-OCT-1999; 990US-0161360P.
 PR 25-OCT-1999; 990US-0161361P.
 PR 25-OCT-1999; 990US-0161404P.
 PR 25-OCT-1999; 990US-0161405P.
 PR 26-OCT-1999; 990US-0161406P.
 PR 26-OCT-1999; 990US-0161359P.
 PR 26-OCT-1999; 990US-0161360P.
 PR 28-OCT-1999; 990US-0161930P.
 PR 28-OCT-1999; 990US-0161932P.
 PR 28-OCT-1999; 990US-0161933P.
 PR 29-OCT-1999; 990US-0162142P.

Query Match 22.7%; Score 538.5; DB 3; Length 823;
 Best local Similarity 33.6%; Pred. No. 4.5-46;
 Matches 153; Conservative 88; Mismatches 178; Indels 37; Gaps 18;

QY 3 KIEVSINSKATINSISTSP-LOPKTCLLQDQTPKAVPIFPPYDIDFHDPLPQTLA 60
 Db 381 OMRYDVSVDLIIKES-SEPPNHLKKFKLSSLEQGPTFEGPMTEFYS-ANNSLKPTPEQ 438

QY 61 DLROALSETLTLLYPLSGRKYNNLYIDFEGVYLERAVNCMTDFFLRLKCEBLNFZV 120
 Db 439 MIKSLSLSLTIIHPLAGLKQKNSIDNDSGAFLERVNSLNLLEPSSDLSLQLI 438

QY 121 PIKFMSMEAISCDERYPLGIVQVMYFDSG-IAIGVSYSHKLIDGGTADCPFLXNGAV-FRG 173
 Db 499 PTSVSDIET---RTLLAQASEFCCGMSIGVCISHLADATSIGFMKSNRAISSRG 554

Qy 179 CRENITHPSLSBAAALLPPP--RDLPEKYYDQMBALWFRGKVAATRFFGYKAISIQ 235
 Db 555 SIKTIGP-VFDIVKCFPPGNSETSPAPVVFPEIM--NQTLSPFPOSSSIQALQ 609
 Qy 236 DEAKSESVPKPSRVAHTGFLKHLIAASRALTSGTSISTRSLIAAQAVNLURMANETVL 295
 Db 610 AKASSFVNQPTVEA7SALIWKSAMKATRVT-SGT--SKPSTLANSVLSRSRSV-PPFT 665
 Qy 296 DNATGNEFWAALELSHTPPISDIKLCDLYNLINGSVKQNGDYFETFKREGYGR- 354
 Db 666 KNSIGNLVSYFAKAE---ESINQKLQYLVSKIRAKQFRDTHIPKLVGNPNATEI 720
 Qy 355 MCEYLDFQRTTSSMEAP--DIYLFLSSWTNF-FNPFLDFGKMGRTSWIGVAGKIESASCKFI 411
 Db 721 ICSE---QKEADMIAQGDFDFWVSAZCRLGLEYTDGKGPVWWGFPSTVRQK--NIV 774
 Qy 412 ILVPTQCGSGTAAWNLLEEKNMAMLEQDPHFLALAS 447
 Db 775 TUDTKEAGGTAAWNLNEQENNLFEQDRELIQFAS 810

RESULT 6
 AAG39694 DE Arabidopsis thaliana protein fragment SEQ ID NO: 49154.
 ID AAG39694 standard; protein; 954 AA.
 XX AC AAG39694;
 XX DT 18-OCT-2000 (First entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 49154.
 XX DE Protein identification: signal transduction pathway: metabolic pathway;
 XX hybridisation assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-00301439.
 XX PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 23-MAR-1999; 99US-0123588P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 01-APR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127452P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0130410P.
 PR 16-APR-1999; 99US-0132048P.
 PR 21-APR-1999; 99US-0133007P.
 PR 23-APR-1999; 99US-0130449P.
 PR 05-MAY-1999; 99US-0132445P.
 PR 06-MAY-1999; 99US-0132446P.
 PR 06-MAY-1999; 99US-0132447P.
 PR 07-MAY-1999; 99US-0132448P.
 PR 11-MAY-1999; 99US-0134246P.
 PR 14-MAY-1999; 99US-0134247P.
 PR 14-MAY-1999; 99US-0134248P.
 PR 14-MAY-1999; 99US-0134249P.
 PR 14-MAY-1999; 99US-0134250P.
 PR 18-MAY-1999; 99US-0134370P.
 PR 19-MAY-1999; 99US-013491P.
 PR 19-MAY-1999; 99US-013491P.

PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146386P.	PR	28-OCT-1999;	99US-0161992P.
PR	02-AUG-1999;	99US-0146386P.	PR	28-OCT-1999;	99US-0161993P.
PR	03-AUG-1999;	99US-0147036P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147204P.			
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	30-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151432P.			
PR	01-SEP-1999;	99US-01519302.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	15-SEP-1999;	99US-0153750P.			
PR	16-SEP-1999;	99US-0154018P.			
PR	22-SEP-1999;	99US-0155130P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	28-SEP-1999;	99US-0156450P.			
PR	29-SEP-1999;	99US-0156595P.			
PR	04-OCT-1999;	99US-0157111P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	13-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0160814P.			
PR	21-OCT-1999;	99US-0160815P.			
PR	22-OCT-1999;	99US-0160981P.			
PR	22-OCT-1999;	99US-0160767P.			
PR	22-OCT-1999;	99US-0160768P.			
PR	25-OCT-1999;	99US-0161404P.			
PR	25-OCT-1999;	99US-0161405P.			
PR	25-OCT-1999;	99US-0161406P.			
PR	26-OCT-1999;	99US-0161359P.			
PR	26-OCT-1999;	99US-0161360P.			
			Query Match	22.7%	Score 538.5 ; DB 3 ; Length 954 ;
			Best Local Similarity	33.6%	Pred. No. 5.7e-46 ; Gaps 18 ;
			Matches	153 ; Conservative	88 ; Mismatches 178 ;
QY		3 KIEVSINSKTHIKPSTSTP--LQPYKTLILDLQTPPVYPIVPTDHFNLNPQTLA 60			
DB	512	QHRYDVVSRDIKPS-STEPMHLKKFKLSSLEQJGPTFQPMYFYS-ANNSIKPTEQI 569			
QY		61 DIOREALSTPDRYPLVQVNNVFDSG-IA1GVSYSKHLID37ADCFLKNSGAV-FRG 178			
DB	630	PTSDVSDT----RTRLJLAQASFFCCSMSICVCSIRKLAQD-SIGFMKSAAISRG 685			
QY		179 CRENLIHPSLSEALLFPF--RDDLPKPYDOMEALPAGSKVATRPFVFGVKAISIQ 235			
DB	686	SIKTKIGAP-VEDIVKIFPGNFSETSPAPVVEPDM----NOTLSKRFIEDSSSIQALQ 740			
QY		236 DBAKSESTPKPSVHATVGFWLWHLIASRALTSGTTSRLLAAQAVNLRTTMNMEVTL 295			
DB	741	AKASSPFENQPTFVEASALIWKSAMQTRTV-SGT-SKEPSLANSLSLRSRVs-PPFT 796			
QY		296 DNATGNYLPTWAAQILESLHTTPRISDLKALCDLWNLNGSVKOCNGDVTETPKGKEGYGR- 354			
DB	797	RNSIGNLQSYFANKAE----EGINQTLVSKIRAKQERDIIHPLKVNPNATEI 851			
QY		355 MCEYLDFFQRTMSSMPEPA--DIXYFSSNTNF-FNPLDGFQWGRTSWIGVAGKIESASCKP1 411			
DB	852	ICSY---QREAGDMLIASQEDDFYTFSSACRFGLYETDGFQWGRPSVQK---NIV 905			
QY		412 1NPTQCCGIEAWVNLEBERMAMLEQPHFLALAS 447			
DB	906	TLLDTKBRAGGIEAWVNBNQEMNLFEQORELLQFAS 941			
		RESULT ?			
		AA2G27611	ID	AA2G27611 standard; protein: 726 AA.	
		AA2G27611;	XX		
		17-OCT-2000 (first entry)	DR		
			XX		
			DE	Arabidopsis thaliana protein fragment SEQ ID NO: 32516.	
			XX		
			AC		
			XX	Protein identification; signal transduction pathway; metabolic pathway;	
			XX	hybridisation assay; genetic mapping; gene expression control; promoter;	
			XX	termination sequence.	
			OS	Arabidopsis thaliana.	
			XX		
			PN	BP1033405-A2.	
			XX	06-SEP-2000.	
			XX	25-FEB-2000; 2000SEP-00301439.	
			XX	25-FEB-1999; 99US-0121825P.	
			PR	05-MAR-1999; 99US-012318P.	
			PR	09-MAR-1999; 99US-012354P.	
			PR	23-MAR-1999; 99US-012579P.	
			PR	25-MAR-1999; 99US-012626P.	
			PR	29-MAR-1999; 99US-012678P.	
			PR	01-APR-1999; 99US-012746P.	
			PR	06-APR-1999; 99US-0128234P.	

PR	08-APR-1999;	99US-0128714P.	PR	19-JUL-1999;	99US-0144333P.
PR	16-APR-1999;	99US-0129845P.	PR	19-JUL-1999;	99US-0144334P.
PR	21-APR-1999;	99US-0130077D.	PR	19-JUL-1999;	99US-0144335P.
PR	23-APR-1999;	99US-0130449P.	PR	20-JUL-1999;	99US-0144332P.
PR	23-APR-1999;	99US-0130510P.	PR	20-JUL-1999;	99US-014484P.
PR	28-APR-1999;	99US-0130891P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0131449P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132407P.	PR	22-JUL-1999;	99US-0145085P.
PR	06-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145089P.
PR	14-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145192P.
PR	11-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145145P.
PR	14-MAY-1999;	99US-01324218P.	PR	23-JUL-1999;	99US-0145218P.
PR	21-MAY-1999;	99US-0134219P.	PR	23-JUL-1999;	99US-0145224P.
PR	24-MAY-1999;	99US-0134221P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	25-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	19-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	23-MAY-1999;	99US-0134256P.	PR	28-JUL-1999;	99US-0145951P.
PR	14-MAY-1999;	99US-0134218P.	PR	02-AUG-1999;	99US-0146386P.
PR	03-JUN-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	04-JUN-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146389P.
PR	07-JUN-1999;	99US-0136212P.	PR	03-AUG-1999;	99US-0147038P.
PR	10-JUN-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0134941P.	PR	04-AUG-1999;	99US-0147302P.
PR	01-JUN-1999;	99US-0135124P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0135752P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	10-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138847P.	PR	09-AUG-1999;	99US-0147935P.
PR	14-JUN-1999;	99US-0139119P.	PR	10-AUG-1999;	99US-0148171P.
PR	16-JUN-1999;	99US-0139452P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139453P.	PR	12-AUG-1999;	99US-0148341P.
PR	17-JUN-1999;	99US-0139452P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139452P.	PR	13-AUG-1999;	99US-0148648P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139462P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139463P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139750P.	PR	26-AUG-1999;	99US-0150844P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151065P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0153700P.
PR	22-JUN-1999;	99US-0139899P.	PR	27-AUG-1999;	99US-0153758P.
PR	23-JUN-1999;	99US-0140353P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0140695P.	PR	16-SEP-1999;	99US-015438P.
PR	02-JUL-1999;	99US-0142055P.	PR	01-SEP-1999;	99US-0151930P.
PR	06-JUL-1999;	99US-0140923P.	PR	07-SEP-1999;	99US-0154779P.
PR	08-JUL-1999;	99US-0142390P.	PR	10-SEP-1999;	99US-015753P.
PR	16-JUL-1999;	99US-0144085P.	PR	23-SEP-1999;	99US-015765P.
PR	16-JUL-1999;	99US-0144086P.	PR	27-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.

PR 23-JUN-1999; 99US-0140353P.
 PR 24-JUN-1999; 99US-0140354P.
 PR 28-JUN-1999; 99US-0140695P.
 PR 29-JUN-1999; 99US-0140823P.
 PR 30-JUN-1999; 99US-0140991P.
 PR 01-JUL-1999; 99US-0141287P.
 PR 02-JUL-1999; 99US-0141642P.
 PR 02-JUL-1999; 99US-014154P.
 PR 06-JUL-1999; 99US-0142055P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144292P.
 PR 19-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 21-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145919P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147028P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 05-AUG-1999; 99US-0147302P.
 PR 06-AUG-1999; 99US-014732P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 09-AUG-1999; 99US-014760P.
 PR 09-AUG-1999; 99US-014793P.
 PR 09-AUG-1999; 99US-014817P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-014856P.
 PR 13-AUG-1999; 99US-014939P.
 PR 16-AUG-1999; 99US-0149684P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149424P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149922P.
 PR 23-AUG-1999; 99US-0149022P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151063P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 30-AUG-1999; 99US-0151303P.

PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-015163P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0158029P.
 PR 29-SEP-1999; 99US-0158322P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159233P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157532P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158322P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159233P.
 PR 13-OCT-1999; 99US-0159244P.
 PR 13-OCT-1999; 99US-015925P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159337P.
 PR 14-OCT-1999; 99US-0159389P.
 PR 18-OCT-1999; 99US-0159548P.
 PR 19-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-016080P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161044P.
 PR 25-OCT-1999; 99US-0161045P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 25-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-016192C.
 PR 28-OCT-1999; 99US-016192P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 22.5t; Score 533.5; DB 3; Length 823;
 Best Local Similarity 33.3t; Prod. No. 1.5e-45; Mi matches 180; Indels 37; Gaps 18;
 Matches 152; Conservative 87; Mi matches 180; Indels 37; Gaps 18;

Qy 3 KLEVSINSRHTIKPSTSTP--LQYKLTLLDQITPAPVPTFVYPTIDHDFNLPTLIA 60
 Db 381 QMRVDDVSRDIKPS-SPPNHLKKFKISLLEQGPTIFGPMVFFYS-ANNSXKPTEQLQ 438

Qy 61 DIRQASSETLTLYPLSGRVKVNLYLIDPEEGVPTY-LEARVNCMTDLIRKTCBLNFRV 120
 Db 439 MURKSLSETLTHYPLAGRKGNTISIDNDGADFLEARVNSPLSSILBPSDSDLQLI 498

Qy 121 PIKPFMSBASIDERYPLIGQVNNEFSG-IATGVSYSHKLIDCGTADCFLKSGAV-FRG 178
 Db 499 PTWSWEST---RTTALLAGA-SFEGCGSMSIGVCLSKLPLAATSGCFLMKSWAISRG 554

Qy 179 CREATIIFSLSEALLPFP--RDLPLPEKYDOMEALWFAGKCKVATRFPVFGYKAISIQQ 235
 Db 555 SIKTIGAP-VEDVYKFFPGNSETSAPVVEDEIM---NOTLSRKFIEDSSSIIQLQ 609

Qy 236 DEAKSESYPKPSRHYAATGELKHLIAASRALTSITSLSTAQAAYNLRTMNMETYL 295
 Db 610 AKASSEFYQNQPTVEAVALIWSAMATRIV-SGT--SKPSLANSASLRSSESV-PPFT 665

Qy 296 DNGATGNLFWWAQARLLESHTTPBISJJKLCDLVNLNGSVKQNCNGDYFETFKGKEGYGR- 354

PR	08-JUN-1999;	990US-0138094P;
PR	10-JUN-1999;	990US-0138540P;
PR	10-JUN-1999;	990US-0138847P;
PR	14-JUN-1999;	990US-0139119P;
PR	16-JUN-1999;	990US-0139452P;
PR	16-JUN-1999;	990US-0139453P;
PR	17-JUN-1999;	990US-0139492P;
PR	18-JUN-1999;	990US-0139455P;
PR	18-JUN-1999;	990US-0139456P;
PR	18-JUN-1999;	990US-0139456P;
PR	18-JUN-1999;	990US-0139457P;
PR	18-JUN-1999;	990US-0139459P;
PR	18-JUN-1999;	990US-0139459P;
PR	18-JUN-1999;	990US-0139460P;
PR	18-JUN-1999;	990US-0139461P;
PR	18-JUN-1999;	990US-0139462P;
PR	18-JUN-1999;	990US-0139463P;
PR	18-JUN-1999;	990US-0139463P;
PR	18-JUN-1999;	990US-0139750P;
PR	18-JUN-1999;	990US-0139763P;
PR	21-JUN-1999;	990US-0139818P;
PR	22-JUN-1999;	990US-0139899P;
PR	23-JUN-1999;	990US-0140353P;
PR	23-JUN-1999;	990US-0140695P;
PR	24-JUN-1999;	990US-0140695P;
PR	28-JUN-1999;	990US-0140823P;
PR	29-JUN-1999;	990US-0140911P;
PR	30-JUN-1999;	990US-0141287P;
PR	01-JUL-1999;	990US-0141842P;
PR	01-JUL-1999;	990US-0142144P;
PR	02-JUL-1999;	990US-0142055P;
PR	06-JUL-1999;	990US-0142390P;
PR	08-JUL-1999;	990US-0142833P;
PR	09-JUL-1999;	990US-0142930P;
PR	12-JUL-1999;	990US-0142977P;
PR	13-JUL-1999;	990US-0143542P;
PR	14-JUL-1999;	990US-0143674P;
PR	15-JUL-1999;	990US-0144005P;
PR	16-JUL-1999;	990US-0144085P;
PR	16-JUL-1999;	990US-0144086P;
PR	19-JUL-1999;	990US-0144252P;
PR	19-JUL-1999;	990US-0144311P;
PR	19-JUL-1999;	990US-0144332P;
PR	19-JUL-1999;	990US-0144333P;
PR	21-JUL-1999;	990US-0144334P;
PR	22-JUL-1999;	990US-0144335P;
PR	20-JUL-1999;	990US-0144352P;
PR	20-JUL-1999;	990US-0144332P;
PR	20-JUL-1999;	990US-0144384P;
PR	21-JUL-1999;	990US-0144516P;
PR	21-JUL-1999;	990US-0145088P;
PR	22-JUL-1999;	990US-0145085P;
PR	22-JUL-1999;	990US-0145087P;
PR	22-JUL-1999;	990US-0145113P;
PR	23-JUL-1999;	990US-0145119P;
PR	23-JUL-1999;	990US-0145145P;
PR	02-AUG-1999;	990US-0145218P;
PR	02-AUG-1999;	990US-0145224P;
PR	26-JUL-1999;	990US-0145276P;
PR	27-JUL-1999;	990US-0145389P;
PR	27-JUL-1999;	990US-014518P;
PR	27-JUL-1999;	990US-0145913P;
PR	03-AUG-1999;	990US-0147038P;
PR	04-AUG-1999;	990US-0147204P;
PR	04-AUG-1999;	990US-0147302P;
PR	05-AUG-1999;	990US-0147192P;
PR	05-AUG-1999;	990US-0147260P;
PR	06-AUG-1999;	990US-0147303P;
PR	06-AUG-1999;	990US-0147303P;

Db	287	LIAESLIGNIM-FSSVVLGTS----QOEELKIEKAVRDLR---KQ-GDDQDNVTKBEGG	336
Qy	352	-----YGENCEYELDFQRTMSMSEBAPDIXLESSWTNPFRL--DFG	391
Db	337	SSAMIGSKLNLMLTNYSKCL-----SYTHEP-----XTVSSWCKL-PLYEASPGW	382
Qy	392	GRTSMIGVAKSIEASACKPFLIVPQCGSSIEAWNLEBKGAMLEQDPFLALAS--PK	449
Db	383	GSPWLI-AGRVAPMILENNWMLIDSRKGQSGIEAFVLPBEMMISQEQTEFLAASVNS	440
Qy	450	TLI 452	
Db	441	VIL 443	
Db	RESULT 12		
Qy	AB911119	AB911119 standard; protein; 436 AA.	
XX			
AC	AB911119;		
XX	31-MAY-2002	(first entry)	
XX	DE	Herbicidally active polypeptide SEQ ID NO 330.	
XX	KW	Herbicidal; Plant; agriculture; herbicide.	
OS	Arabidopsis thaliana.		
XX	W0200210210-A2.		
XX	PD	07-FEB-2002.	
XX	PF	28-AUG-2001; 2001WO-EPO09892.	
XX	PR	28-AUG-2001; 2001WO-EPO09892.	
XX	PA	(PARB) BAYER AG.	
XX	PI	Tietjer K, Weidler M;	
XX	DR	WPI: 2002-269010/31.	
XX	PT	Identifying plant target proteins for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from non-plant organisms.	
XX	PT	PT from plant with nucleic acid or amino acid sequences from non-plant organisms.	
XX	PS	SEQ ID NO 330; 261PP + Sequence Listing; English.	
XX	CC	The invention relates to identifying target proteins (ABB90790-ABB94016) for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for herbicides. The identified modulators are useful as herbicides	
XX	CC	Sequence 436 AA;	
Qy	Sequence 436 AA;		
Query Match	21.0%	Score 498; DB 5; Length 436;	
Best Local Similarity	30.9%	Pred. No. 2.7e-42;	
Matches	145;	Mismatches 184;	
Qy	1	MECKEVSINSKHTIKPSTSSTP--LQPKXKLTLDQLTTPAYVPIFVPTDHD--FNL 55	
Db	1	MEK-NVEILSRLEIVKPS-SPTPDKRKLNLSSMTCALLFYAAFPNLLGFS 58	
Qy	56	PQTLADLQALSETUTLYPLPSGRVKNLYLDDREEGYPYLEARVNCDMTDFLRLKTEC 115	
Db	59	EETSLK-KKSLSKTLPYPLAGRLIGS-FVCECNDEGAVFTEARYDHILSEFLKCPVFS 117	
Qy	116	INEFVPIKPPSMEAISDERYPULGTYQWVFD-SGIAIGVSYSHKLIDGGTADCFLKSWGA 174	
Db	118	LELLIVEAKREAVT--WVLLLIGANFFSGGLVITICVSHKTTDATSLAMFIRGWA 174	
Qy	175	VFRGCEENIHPSLSSRAALLEPPRDLEPKVQDMALWFGRKVTTERFVGRASSI 234	
Db	175	SSRGCTTLI-PSFIASEVFKPDLPEPSKEMDRKREV--FEMSCVTKRFVDBSCKKL 231	
Qy	235	QDEARSESVPPRSPRVAEVATGFLWKLHIAASRALTSCTTSTRSIAQAVNTRMNEY 294	
Db	232	RAKASNNLVXPKPTVREAVTALFWRCUTKVSEL---SSLTPRTSVLQILVNLRGK-VDSL 286	
Qy	295	LDNATGNLFWW----AQATLELSHTTPESDLKLCDLVNLNGSVKQNGDVFETEK- 347	
Db	287	CENTIGNMLSMILKNEEALEIPIOVDEIRRAK-EIISLNCKEMSSKSSRIPFLEE 344	
Qy	348	-GKEGYGRMCEYLDQFTQTMSSMPEADDYFESSWTF-FNPFLDFGMGRTSWIGVAKSIES 405	
Db	345	IGKV-YGRGNEM-----DLWNSNSNCKLGYDADFGWKPWV-WTGRGTS 387	
Qy	406	ASCKETLILVPTQCGSGJEAVNLEEKMANLEQDDPHFLAS--PKTLI 452	
Db	388	HFKNLMLLIDTQDKDGEIGEAWITLREQMSLFECQDQELLESASLNPVLI 436	
RESULT 13			
Qy	AD91973		
XX	ID	AD91973 standard; protein; 436 AA.	
XX	AC	AD91973;	
XX	DT	04-DEC-2003 (first entry)	
XX	DE	Acyltransacylase enzyme #11.	
XX	KW	Acyltransacylase; transacylase; paclitaxel; taxoid; paclitaxel biosynthetic pathway; enzyme.	
XX	OS	Arabidopsis thaliana.	
XX	XX		
XX	XX	US2003106891-A1.	
XX	XX	US2003106891-A1.	
XX	XX	12-JUN-2003.	
XX	XX	18-SEP-2002; 2002US-00166984.	
XX	XX	PR 30-SEP-1999; 99US-00411145.	
XX	XX	PR 07-DEC-1999; 99US-00457046.	
XX	XX	PR 25-MAY-2001; 2001US-00866570.	
XX	XX	(UNIW) UNIV WASHINGTON STATE RES FOUND.	
XX	XX	PA Croteau RB, Walker KD, Schoendorf A, Wildung MR;	
XX	XX	DR WPT; 2003-659103/62.	
XX	XX	New purified protein useful for producing paclitaxel and related taxoids.	
XX	XX	Example 1; Fig 6; 135PP; English.	
CC	CC	The invention relates to a transacylase enzyme and the polynucleotide encoding it. The invention also relates to a recombinant nucleic acid molecule comprising a promoter sequence operably linked to the polynucleotide encoding the transacylase enzyme and a cell transformed with the recombinant nucleic acid. The transacylase enzyme is useful for producing paclitaxel, related taxoids and intermediates in the paclitaxel biosynthetic pathway. This sequence represents an acyltransacylase enzyme of the invention.	
CC	CC	Sequence 436 AA;	
PS	PS	XX	
Query Match	21.0%	Score 498; DB 5; Length 436;	
Best Local Similarity	30.9%	Pred. No. 2.7e-42;	
Matches	145;	Mismatches 184;	
Qy	1	MECKEVSINSKHTIKPSTSSTP--LQPKXKLTLDQLTTPAYVPIFVPTDHD--FNL 55	
Db	1	MEK-NVEILSRLEIVKPS-SPTPDKRKLNLSSMTCALLFYAAFPNLLGFS 58	
Qy	56	PQTLADLQALSETUTLYPLPSGRVKNLYLDDREEGYPYLEARVNCDMTDFLRLKTEC 115	
Db	59	EETSLK-KKSLSKTLPYPLAGRLIGS-FVCECNDEGAVFTEARYDHILSEFLKCPVFS 117	
Query Match	21.0%	Score 498; DB 7; Length 436;	

Best Local Similarity 30.9%; Pred. No. 2.7e-42; Matches 145; Conservative 90; Mismatches 184; Indels 50; Gaps 19; Qy 1 MEKIEVSINSCHTIRPSSTSSP--LQYKUTLIDQUTPPATVPIVFFYPTIDHD--FNL 55 Db 1 NEK-AVEILSERKPS-SPPDDKRLINLSSLDIISSPMTGALLFYAAPQNLIGEST 58 Qy 56 FQTLADLROALSETLTIYPPSGRVKNLYDDEEGVPIEARVNCOMDTFLRKIBC 115 Db 59 BETSLKIKLSSKLTLPYPPAGRIGTS-FVCECDGAEVFAEYDDEEGVPIEARVNCOMDTFLRKIBC 117 Qy 116 LNEFVDPKPFSEMAISDERYPLLGDVNVFED-SGIAIGVSYSHKLIDGTTACDFLKSGWGA 174 Db 118 LELLIPVEAKSREAVT--WVYLLQIANFRCGGLVITICSHKLTIDATSLAMEFRGWA 174 Qy 175 VFRGERENITHPSLSEALLIPERDOLPCEYDQMEALWPAKKVYATRTPFGVKAISI 234 Db 175 SRRGSGITL-PSPTASBVFVFKPDLDEPSKMDRESEV--EEMSCVTRKTFEDASKIKKL 231 Qy 235 QDEAKSESPVSPRSVHATVGPFLWKHLIAASSEALTSCTTSRSLTIAQAVNURTMAMMETV 294 Db 232 RAKRSRNLVNPRTVHATVAFWRCVTKYSLR--SSLTPRTSVLQILVNLGRK--VDSL 286 Qy 295 LDNATGNLFWW---AQNLLELSHTTPESDLKLCDLWNLINSSVKQNCNGDYPBTFK- 347 Db 287 CENTIGNMILSMILNEEAATERIQDWDVDEERRAK--ELFSLNCKEMSSSSRIFELLE 344 Qy 348 -GREGYGRMCEYLDQFQTMSSMEPADTIVLSSWTNF-EPNPLDGFGRITSGVAGKIES 405 Db 345 IGKV-YGRNEM-----DLWNNSWCKGLYDADFGKPYW--VTGRGTS 387 Qy 406 ASCKFLVLPVQCGSSIEANVNLNEEKMAMEQDPHFLALAS--PKTLL 452 Db 388 HFKNMLLIDTKDGEIAWITLFBQMSLFECDQELLESAISNPPVLI 436

RESULT 14
ID ADB91985 standard; protein: 446 AA.
XX ADB91985;
XX 04-DEC-2003 (first entry)
DE Acyltransacylase enzyme. #23.
XX Acyltransacylase; transacylase; paclitaxel; taxoid;
KW paclitaxel biosynthetic pathway; enzyme.
XX Arabidopsis thaliana.
OS

US2003108891-A1.

XX 12-JUN-2003.

XX 18-SEP-2002; 2002JUS-00166984.

XX 30-SEP-1999; 99US-00411145.

XX 07-DEC-1999; 99US-00457046.

XX 25-MAY-2001; 2001US-00866570.

XX (UNIW) UNIV WASHINGTON STATE RES FOUND.

XX Croteau RB, Walker KD, Schoendorf A, Wildung MR;

XX DR 2003-559103/62.

XX PT New purified protein useful for producing paclitaxel and related taxoids.

XX PS Example 1; Fig 6; 135pp; English.

XX CC The invention relates to a transacylase enzyme and the polynucleotide

CC encoding it. The invention also relates to a recombinant nucleic acid

CC molecule comprising a promoter sequence operably linked to the
CC polynucleotide encoding the transacylase enzyme and a cell transformed
CC with the recombinant nucleic acid. The transacylase enzyme is useful for
CC producing paclitaxel, related taxoids and intermediates in the paclitaxel
CC biosynthetic pathway. This sequence represents an acyltransacylase enzyme
XX of the invention.

Sequence 446 AA;
Query Match: 20.4%; Score 485; DB 7; Length 446;
Best Local Similarity 29.1%; Pred. No. 6.2e-41;
Matches 143; Conservative 91; Mismatches 165; Indels 92; Gaps 23;
Qy 4 IEVSINSKHTIKPSTSSTP--LQYKLTIDLTQDLPPTAYVPIVFFYPTIDHD-ENFLCTLA 60
Db 6 MKVETISKEIILKPS-SPTPNNTQTCIOTLQSTIDHILLPPVYTAFLY--TKNNDLISCFHTSH 62
Qy 61 DLQQLASSTLTLYPLSGRVKNLYDDEEGVPIVARY-NCOMDTFLRKIECLNEF 119
Db 63 KLTTSUSETLTKFYPVGRIL-TGTVTDCDTEGAIFLARYVNCNPTEFLCPDFDAAQQL 121
Qy 120 VPKDFSMEAISDERYPLLGQVQVNF-SGIAIGVSYSHKLIDGTTACDFLKSGWAFRG 178
Db 122 LPDLYDNDPVTAAATWPLLLVYKATYFGGGMAGIGCITHKIAADASLSTFRSWATARG 181
Qy 179 CREN---IIHPSLSEALLIPPRDDEKPEKVDQMBALWFGKRYA-TERRFVGKAISSI 234
Db 182 -ENDAAAMESPVFGANFPVPAANFPVPAKLUPADEQ---AGKRSSTTKEVFEASKVIEDL 234
Qy 235 QDEAKS-ESVPKPSRVHATGFLWKLHIAASRLSTGTSSTRLSAAQATNLRTENM-MET 293
Db 235 RTKQASEETDQDPTKVESTALIWKCFVASSK-----TTCDHKLVLQIANLRSCL-PSL 288
Qy 294 VLDNATGILFWWA-----CAILELSHTTPPEISLKLCD-----326
Db 289 LQESIGNLMESSSVLISGRGGEYKIEAVDRJRKKEELGTVILDEGGSSDSSSMIGSK 348
Qy 327 LVNLNGSTKQCNEDYFETFKGKEGYGRMCEYLDQRTMSSMEPADTIVLSSWTNFNFP 386
Db 349 LANLM-----LTNYSLR-----SYTHEP---PTVSSWCKL--P 377
Qy 387 L---DEGWGRTSWIGVAKTIESASCKFLILVPTQGSGIEAANVLEERKAMLEODPHFL 443
Db 378 LYEAESFGWDSPVW-VVGNVSPVGLNAMLIDSQDGQGEAVTLPPEMNNSFFQNPELL 435
Qy 444 ALA--SPKTLI 452
Db 436 AFATMNPSPVLF 446

RESULT 15
ID ADB91986 standard; protein: 435 AA.
XX ADB91986;
AC ADB91986;
XX 04-DEC-2003 (first entry)
DE Acyltransacylase enzyme #24.
XX Acyltransacylase; transacylase; paclitaxel; taxoid;
KW paclitaxel biosynthetic pathway; enzyme.
XX Arabidopsis thaliana.
OS

XX 18-SEP-2002; 2002JUS-00166984.

XX 30-SEP-1999; 99US-00411145.

XX 07-DEC-1999; 99US-00457046.

XX 25-MAY-2001; 2001US-00866570.

XX (UNIW) UNIV WASHINGTON STATE RES FOUND.

XX Croteau RB, Walker KD, Schoendorf A, Wildung MR;

XX DR 2003-559103/62.

XX PT New purified protein useful for producing paclitaxel and related taxoids.

XX PS Example 1; Fig 6; 135pp; English.

XX CC The invention relates to a transacylase enzyme and the polynucleotide

CC encoding it. The invention also relates to a recombinant nucleic acid

Search completed: July 6, 2004, 13:36:43
Search time: 60.6787 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	498	21.0	436	3	US-09-457-046B-62
2	485	20.4	446	3	US-09-457-046B-74
3	448	18.9	435	3	US-09-457-046B-60
4	404.5	17.0	433	3	US-09-457-046B-66
5	387.5	16.3	439	3	US-09-457-046B-68
6	308	13.0	331	3	US-09-457-046B-59
7	242	10.2	461	1	US-09-457-046B-64
8	230	9.7	448	1	US-08-207-904-2
9	230	9.7	448	1	US-08-207-904-17
10	228	9.6	451	3	US-09-457-046B-69
11	225	9.5	445	3	US-09-457-046B-73
12	222	9.4	497	3	US-09-457-046B-65
13	219	9.2	482	3	US-09-457-046B-63
14	179	7.5	443	3	US-09-457-046B-50
15	178.5	7.5	461	3	US-09-457-046B-70
16	177.5	7.5	441	3	US-09-457-046B-54
17	177.5	7.5	458	3	US-09-457-046B-61
18	176.5	7.4	306	3	US-09-457-046B-22
19	175	7.4	455	3	US-09-457-046B-72
20	173.5	7.3	460	3	US-09-457-046B-71
21	157	6.6	438	3	US-09-457-046B-58
22	154.5	6.5	450	3	US-09-457-046B-67
23	150	6.3	303	3	US-09-457-046B-20
24	148.5	6.3	440	3	US-09-457-046B-45
25	146.5	6.2	459	4	US-10-074-279-2
26	146.5	6.2	459	4	US-10-074-279-2
27	144	6.1	439	3	US-09-457-046B-28

Qy 406 ASCKPILILVPTQCSSGIEAVNLEBEMAMLEQDPFLALAS--PRTLLI 452
 Db 388 HEKNUMLLIDTKDSEGIEAMTILIBEQMSLFEQDQELLESASLNPVLL 436

RESULT 2
 ; Sequence 74, Application US/09457046B
 ; Patent No. 628735
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney et al.
 ; TITLE OF INVENTION: Transacylates of the Paclitaxel Biosynthetic Pathway
 ; FILE REFERENCE: 53679
 ; CURRENT APPLICATION NUMBER: US/09/457,046B
 ; CURRENT FILING DATE: 1999-12-07
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 74
 ; LENGTH: 446
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-457-046B-74

Query Match 20.4%; Score 485; DB 3; Length 446;
 Best Local Similarity 29.1%; Pred. No. 7.3e-44; Gaps 23;
 Matches 143; Conservative 91; Mismatches 165; Indels 92; Gaps 23;

Qy 4 LEVINSKHTIKPSTSSTP--LQPYKLTLDQTPPAYPPIVFFYPTDHD-FNLPTOLA 60
 Db 6 MARYTISKEIKPSS-SPTPNNLQLTSYDHLPPVTVAFLY--TNDLISQERTSH 62
 Qy 61 DILROALSETLTPLSGRVKNNYIDDEEGFVPLAEV-NCMDTDFRLRKECLEF 119
 Db 63 KLTKLSETLTKEYPLAGL-T367VDCDUGALFVDAVNNCPLTEFLKCPDDEALQOL 121
 Qy 120 VPIKPKFMSAISDERYPLIGVQARVFD-SGIAIGVSVSHKLIDGTDGFLKNSGAVFRG 178
 Db 122 LFLDVNDVNPYVAATWPLLIVKATYFGCCGMAIGCITHKADASISTPISWAATARG 181
 Qy 179 CREN--LTHPSLSEAALLFPPRDLPEXVQDOMEALWFAGKIIA-TRRFVGKAISI 234
 Db 182 -ENDAAAMBSPVAGANFYPANBAKLPDEQ---AGROSSITKRFVFEASKVEDL 234
 Qy 235 QDEAKS-BSVPKPSRVAHTVGFMLKHTAASRAUTSGTTSLSIAQAVNLSTRANNET 293
 Db 235 RTRKASEETVQDPTRVESTVTLWRCFVASSK---TTCDHRYVQLANLRSKI-PSL 288
 Qy 294 VLDNATGNLFWIA-----QALELSEHTTPBISDLKLCD----- 326
 Db 289 LOBSSIGNLMFSSVTLSIGREGEVKIEAVDLRKKKBELLGTVLDEGSSDSSMIGSK 348
 Qy 327 LWNLLNGSYTKQNCNDYFETFKGEGYGRNCYLDQFRTMSMSEPAFDIYLFSSTWNFPNP 386
 Db 349 LANIM-----LTNYSRLL-----SVEHEP---YTVSSNCKL--P 377
 Qy 387 L---DEGMGRTSNIGVAGKIBEASCKTFLINTPTQCGS31EAWNLEBEMAMLEQDPFL 443
 Db 378 LYRASPGMDSPVW--VVGNVSPVGNLAMLDSDKGQTEAFTLPEENMASPFONPEL 435
 Qy 444 ALA-SPKTLL 452
 Db 436 AFATMNEPSVLV 446

RESULT 3
 ; Sequence 60, Application US/09457046B
 ; Patent No. 628735
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney et al.
 ; TITLE OF INVENTION: Transacylates of the Paclitaxel Biosynthetic Pathway
 ; FILE REFERENCE: 53679
 ; CURRENT APPLICATION NUMBER: US/09/457,046B
 ; CURRENT FILING DATE: 1999-12-07
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 60
 ; LENGTH: 435
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-457-046B-60

Query Match 18.9%; Score 448; DB 3; Length 435;
 Best Local Similarity 29.8%; Pred. No. 7.5e-40; Gaps 15;
 Matches 137; Conservative 79; Mismatches 193; Indels 50; Gaps 15;

Qy 3 KIEVSINSKHTIKPSTSSTPLOPKLTLDQTPPAYPPIVFFYPTDHD-FNLPTOLA 62
 Db 4 KLEVTT--GKEVKEP-SPSPDRLLSILLYCPGIVSTIFFDLTESSEV- FSENLL 58
 Qy 63 RQALSETLTLYPLSGRVKNNYIDDEEGFVPLAEVNCMDTDFRLRKIECLNEFVPI 122
 Db 59 RKLISBTLTSFYPYLAGIE-GLSSISCNDEAVFTARTDILPDLRNLNTDSLGFLPT 117
 Qy 123 KPFMSMEAISDERYPLIGVQANVFP--DGSIAIGVSYSHKLIDGTDGFLKNSGAVFRGCR 180
 Db 118 --LAAGEPSPAAPWPLSTKVTFPGSGSYAVSVSYSHKICDIASTVTFYDWAATTAKGK 174
 Qy 181 EN-----1HPSLSEALLFPPRD---DLBPKYDOMEALWFAGKVKATRFFVFGV 228
 Db 175 SNTSIEFAETTIYPPPSHMYBQPSTSDSNITSKVV-----LKFVFP 220
 Qy 229 RAISIQQDEAKSESVKPSRVAHTVGFMLKHTAASRALTSGFTSTRSLAAQAVNLRT 288
 Db 221 SKALBLKKAASESVPKPSRVAHTVGFMLKHTAASRALTSGFTSTRSLAAQAVNLRT 277
 Qy 289 MNVETVLDDNATGNLFWIAQALELSEHTTPBISDLKLCDLVNLNGSVQNCNDYFETFKG 348
 Db 278 IPSSVAPRQVIGN---QSGFSLKKDAB--SEREPIPBIVATPKNKRKVNEM-LKESLQG 331
 Qy 349 KEGYGRMCEXYLDQFTMSMSEPAFDIYLFSSTWNFPNPUDFGGRTSWIGVAGKIESAS 407
 Db 332 NT-1GQSLSLMAETVSESTE-IDRYSMSWCRKPFYEDFGSSPPWWGYASHTIYDN 388
 Qy 408 CKFTILVPTOCGSGIEAWWNLREEKMAMLEQDPFLALA 446
 Db 389 MVGVVLIDSKEGDTTEAMSLPPEEDMSVYFDDQELLAYA 427

RESULT 4
 ; Sequence 66, Application US/09457046B
 ; Patent No. 628735
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney et al.
 ; TITLE OF INVENTION: Trinitacylates of the Paclitaxel Biosynthetic Pathway
 ; FILE REFERENCE: 53679
 ; CURRENT APPLICATION NUMBER: US/09/457,046B
 ; CURRENT FILING DATE: 1999-12-07
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 66
 ; LENGTH: 433
 ; TYPE: PRT
 ; ORGANISM: Clarkia breweri
 US-09-457-046B-66

Query Match 17.0%; Score 404.5; DB 3; Length 433;
 Best Local Similarity 28.0%; Pred. No. 4.1e-35; Gaps 21;
 Matches 137; Conservative 80; Mismatches 171; Indels 101; Gaps 21;

Qy 4 IEVSINSKHTIKPSTSSTP--LQPYKLTLDQTPPAYPPIVFFYPTDHD-FNLPTOLA 61
 Db 1 MNVTMHSKLLKPKSTI-PTPNHLQKLNLSLSDQIQIPFYGLIFFYETSLSDNSDI-TLSK 57
 Qy 62 LRQALSETLTLYPLSGRVK-NNLXLYDFFEGVPLAEVNCMDTDFRLRKIECLNEFV 120

SEQ ID NO 64
LENGTH: 461
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-09-457-046B-64

Query Match 10-2%; Score 242; DB 3; Length 461;
Best Local Similarity 25.4%; Pred. No. 2.2e-17; Indels 108; Gaps 27;
Matches 120; Conservative 76; Mismatches 169; Gaps 27;

Qy 4 LEVSINSKHTIKPSTSSTPLQP-----YKLLDQLTPPAYVPIVFFPTIDHDNLQ 57
Db 15 IPIVTINGQFLVHPS-SPTPANOSPHHSVLNSLNDLIGARVFTSVYFPTSTNNRESP-- 71

Qy 58 TLLADROALSTLTLVYPLSPGRVK--NNNTYDDE--ECPVYLEARVYCDMIDFLRLK 112
Db 72 VLRQLDASEVLPVPPYPLSLRERVNGKLEVEFGEBGVLVMSANSMDLAD----- 125

Qy 113 IECLENEFPIKPKPSMEAI----SDERY----PILGVYQVNFD-SGIAIGVSVSHKLIDG 162
Db 126 ---LGDLTVNPANDELLPRNGPBEAYKILLEMPILLIAQVTFPRGGFSIGRLHCICG 182

Qy 163 GTADCEFLKSGWAVPRGCRENIIHPS-----LSRAMELFPPRDOLP-EKVDOMEA 211
Db 183 FGMQFGLSWATAK-TGKLLADEPVDRETFXPNNPDMVYFEEHXYLPIEERSNLNTS 241

Qy 212 LWPGKVK-----ATTRFPGKRIASSITDDEAKSESVXPSPR-----HATVGFJWKLIA 262
Db 242 LMOT-KPLOKCYNISKEFQCRVSIACGD-----PTLVCSTEDAMAATIWRSMV- 290

Qy 263 ASRALTSGCTSTRSLIAAGAVNTRTRMMNETVLDNATGNLFWAQAIL-----ELSH 314
Db 291 --KALDVKPLDNNLRLT-SVWTRTRLTLKLRKGPIGRVWTCACMSVSELINDSLSK 347

Qy 315 RTPEISDLKICLQDLYNLINLGSVKQCGNDYFTFKSGKEGYGRMCAYLDFORTMSSMEPADI 374
Db 348 TTRIVQDARL-----RVSEDTLRS-----MOTYDVRP-KRLEFGKL 385

Qy 375 YLSSWNTW--FNUFLDGSGRTSNGVACKIE-SASCKFLILPFTQCGSSGIE 424
Db 386 TI-TOWTRFEMVYTADEFGNGKPVY--AGSPIDL.RPTPQCVLFLQ--GGVSE 431

RESULT 8
US-09-207-904-2
Sequence 2, Application US/08207904
GENERAL INFORMATION:
APPLICANT: Tuttle, AnnMarie
APPLICANT: Crossland, Lyle D.
TITLE OF INVENTION: Another-Specific cDNA Sequences, Genomic DNA Sequences and Recombinant DNA Sequences
TITLE OF INVENTION: Another-Specific cDNA Sequences, Genomic DNA Sequences and Recombinant DNA Sequences
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
CITY: Hawthorne
STATE: New York
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/207,904
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/908,242
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: CGC 1624
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8615
TELEFAX: (919)541-8659
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 448 amino acids
TYPE: amino acid
TORSIOL: linear
MOLECULE TYPE: protein
US-08-207-904-2

Query Match 9.7%; Score 230; DB 1; Length 448;
Best Local Similarity 23.1%; Pred. No. 4.3e-16;
Matches 107; Conservative 76; Mismatches 193; Indels 88; Gaps 20;

Qy 4 IEVSINTSKHTIKPSTSSTPLQPYKUTLQLTQPAYVPIVFFPTIDHDNLQOT--LA 60
Db 1 MKEVLSKHHVVKPA-EATWNGTVSLSECQTFATHVPTIYTFRC-HDC-LPSTDNIK 57

Qy 61 DLQALSELTTLTYPPLSPLRKV-----NNLYTDDPREGVPLLEARVNCMDMDFLRLRKIECL 116
Db 58 TLRLSSLKALVHFPPLSERLWRAGSLDNCNASGIVLMEAEETAKLDDLGDFSESPDIL 117

Qy 117 NEFPIKPKPSMEAASDERYPLLQYQVNFDD-SGIAIGSVSHKLIDGADCFELKSGAV 175
Db 118 NSLPRVDTY--IPIDELPLFLVQLTQPGGITALSFISHAWDQGQALYFLETEWASL 174

Qy 176 FRGCENTIHPSLSEAAFLFPPRDLPLPCKVDPOMEALWPAKRVATRPFVGKAISSTQ 235
Db 175 ARG--EPLNEPFDIKF--RAGEPAPYPTPEHLQNPPLQG-----SSSE 221

Qy 236 DEAKSESYVKPSRVRHATVCPFLKHLIASRALTS--GTTSPRLSIA----- 280
Db 222 EKKEONET--KGSMILK-----LTGHQVEMURKKAQNGNQERSYTRYEVYAHWRACKAR 274

Qy 281 -----QAVNLTRRNNMETYVLDNATGFLWQAELL-----SHITPEISDLRCLDVLNLNG 333
Db 275 GHKFQQPNTLICVNTRIMQPLPKSY-FGNAIIVDVAANGYSSDITSRPLEYARRVIA 333

Qy 334 SVKCGCNGDYFETFKGKBGYGRMCAYLDFORTMSSMEPADIYFLSSWTFEE--NP---- 386
Db 334 AIRVUTSYDANS-----TIDLNQEDLSKQDIAHFRSKS23PFPYKPLAVI 381

Qy 387 -----LDFGWRGRTSWMIGAKIESASCKFLILVPTOCGSG 421
Db 382 SWTSPLPLQDGFMGKETHMSPGTHEYDGC-----VILPGKEGDD 422

RESULT 9
US-08-207-904-17
Sequence 17, Application US/08207904
GENERAL INFORMATION:
PATENT NO. 5477002
GENERAL INFORMATION:
APPLICANT: Tuttle, AnnMarie
APPLICANT: Crossland, Lyle D.
TITLE OF INVENTION: Another-Specific cDNA Sequences, Genomic DNA Sequences and Recombinant DNA Sequences
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
CITY: Hawthorne
STATE: New York
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/207,904
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/908,242
FILING DATE:
ATTORNEY/AGENT INFORMATION:

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.C, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/207,904
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/908,242
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Lazar, Steven R.
 REGISTRATION NUMBER: 32,618
 REFERENCE/DOCKET NUMBER: CGC 1624
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (919)541-8615
 TELEFAX: (919)541-8689
 INFORMATION FOR SEQ ID NO: 17:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 448 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-207-904-17

Query Match 9.7%; Score 230; DB 1; Length 448;

Best Local Similarity 23.1%; Pred. No. 4.3e-16;

Matches 107; Conservative 76; Mismatches 193; Indels 98; Gaps 20;

Query 4 IEFINSKHTIKPSTSSPLQPKLILLDQITPAVYVPPFPIITDFFNPIQTC--LA 60

Db 1 MKVSLKHEHVKVPA-EATWNGTYSLSCEQTFAVTHVPPYYRRC-HDC-LESTDNTIK 57

Qy 61 DLRAALSETLTLTYPPIGRK--NNLYIDDFEEGVPPVLEARNVCDMTDLIRRKICL 116

Db 58 TIKTSLSKAVHPPFPLSBRPLWAGSRPLDCCASGIVTMRAEAKADDLGDFSPSPBDI 117

Qy 117 NEFVPIKXPSSMEAISDERYPLLGQVNYED--SGTAIGVSVSHKLIDGTTAFCFLKSAGAV 175

Db 118 NSLSPRVDI--IPIDEPLLFLVLTREQQGGTALSFAISHAVDGSALYFTEWASL 174

Qy 176 TPGCRENTHPSLSEAAI1FPRDPLPEKXVQDMEALMPAKCYATRPEFGYKAISSTQ 235

Db 175 ARG-EPGNEPFDKFL--RAGEPPIAYTFEHLQNPPLLGQ--SSSE 221

Qy 236 DEAKSESPVPSRVRHATVGPFLKHLIAASRALTSS--GTTSPRLSIA-- 280

Db 222 EERKNET-KGSMLK--LTHQVEMURKKANQGNOQSYRVEVTAHWRACKAR 274

Qy 281 ---QAVNLTRRNMETYLTDATGKLNFLWAQATI--SHTTBISDLRKLCDVNLNG 333

Db 275 GHRPQOPTNLICVNRINMOPPLPKSYF-GNNAVDVIANGVSSDITSRPLEYVARRTA 333

Qy 334 SVKCGNDUYFETPKGKEGYGRMCYLYDF-PTMSSMPEADDIYLFSSWTNMF--NP---- 386

Db 334 AIRKNTSDYANS-----TIDEFLRNOEDLSKYYDIHAFRSKEGFYGRNPNLGYTI 381

Qy 387 -----LDFGNGRTSMWGVACKIESACKFLVPPQCGSG 421

Db 382 SWISLPLGPGKKELEMSPGTHEYDGC--VILPCEGDD 422

Qy 457-046B-73

RESULT 11 US-08-457-046B-73

; Sequence 69, Application US/09457046B

; Patent No. 6287835

; GENERAL INFORMATION:

; APPLICANT: Croteau, Rodney et al.

; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway

; FILE REFERENCE: 5379

; CURRENT FILING DATE: 1999-12-07

; NUMBER OF SEQ ID NOS: 74

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO: 73

; LENGTH: 445

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

; US-09-457-046B-73

; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-457-046B-69

Query Match 9.6%; Score 228.5; DB 3; Length 451;

Best Local Similarity 22.6%; Pred. No. 6.4e-16;
 Matches 99; Conservative 70; Mismatches 182; Indels 87; Gaps 19;

Qy 13 TIKESTSSTPLQP---XKLILDQITPPAYDVFIVEYPIITDENDYIQTLLADIRQALAS 67

Db 5 TPERSYTTVBAEPYDWSGREPLAENDQVGTTIPLTYFDKPSSESFO-GNVEVETLKTLSLS 63

Qy 68 ETIILYTPISGRV--KNNLYIDDFEEGVPPVLEARNVCDMTDLIRKICLNBVPVPIK 123

Db 64 RVIYHYPYPMGRALWMLPGRFRENCAGVETIESEGKFLSDF--KDESPTRP 114

Qy 124 PFS--MEALS---DERYPLLGQVYNVFD-SGTAIGVSVSHKLIDGGTACDFCLFLKSAGAVF 176

Db 115 ERENIMPQNYKNPLETEIPLFLAQTKFCGGGSLISVNSHAIYDQSAHLISERGRLA 174

Qy 177 RGCRENTHPSLSEA-----LFPPI-----DDLPKCYDQMEALWFAEKKVATR 222

Db 175 RGEPELTV-PFLDKXILWAGEPLEPFVSPKFDKEFDOPPELGETDNEBKKRTIVY 233

Qy 223 RFFVFGWKAIISSIDBEAKS3SVPVKE--SRVHATVGFMLKHLIAASRALTSGTTSRSLI 278

Db 234 MLFPLSTSQQLKREKANGSKHSDAKFPTGKQYHWR--CACKQARHGSPEQTAALGI 291

Qy 279 AAQAVNLTRRNMETVLNDATGNEFLWAQATI--SHTTPEISDLKCDLVNLNGSVK 336

Db 292 C----IDTRSRMPEPLPPIG---YFENATLQWASTSGELISNELGFAASLISKAK 341

Qy 337 QCNGDYYFETFKGKGGYGRMCYLYDFQRTMSSMPEADDIYLFSSWTNFF--NP---- 386

Db 342 NVTNEYMT-----G-EBYLNQDKDJKFO---DHALGSTERGPYGNPKLGVVSWL 389

Qy 387 -----LDFGNGRTSMWGVACKIESACKFLVPPQCGSG 398

Db 390 TLPMYGLDOWGKEFYTG 407

Qy 457-046B-73

RESULT 11 US-08-457-046B-73

; Sequence 73, Application US/09457046B

; GENERAL INFORMATION:

; APPLICANT: Croteau, Rodney et al.

; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway

; FILE REFERENCE: 53679

; CURRENT FILING DATE: 1999-12-07

; NUMBER OF SEQ ID NOS: 74

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO: 73

; LENGTH: 445

; TYPE: PRT

; ORGANISM: Arabidopsis thaliana

; US-09-457-046B-73

Query Match 9.5%; Score 225; DB 3; Length 445;

Best Local Similarity 22.7%; Pred. No. 1.5e-15;
 Matches 106; Conservative 86; Mismatches 181; Indels 94; Gaps 25;

Qy 4 IEVSINSKHTIKPSTSTPLQPKLTDPLQYKLTLLDOL-TPPAPYPIVFTYPIPTCH----- 51

Db 1 MSIQIKQSTMVRPA-BTPKPSLMSLNIDMILRTYSHTGAVLYKQPDNNDNPHSSS 59

Qy 52 --DENLPQTLADLRQALSETLTYPLSRSRQN--LYIDDPFREGVPLBARYNCDMT 105

Db 60 MYFDANI-----LIALSALKALVPPYPMAGRILKINGDRIDCNABGALFVEAESSHVL 113

Qy 106 DFLRLRKIECLNE-FVPPIKPSMBAISDERYPLLGQVQNVFD-SCIAIGVSVSHKLIDGG 163

; LENGTH: 451

Db	114	DFGDFRNPDELLRVMVPTCDYS-KGISS--FPILMVQLTRRCGGSIGFAQHHHVDCGM	170	Qy	375	---YLFFSSWTFN-FNPUDFGWG-----RTSWIGHAGKIESASCKFLILVPTQCGSGI	422		
Qy	164	TADCFLKSGVAFRG-----PSLESEAALLFPP-RDDIPEKYUDQ	209	Db	413	DGPAFTVSSGRSPVNOVDFEGNSPFFSSYHFIFWGGSSAG-----YVMPMP---	460		
Db	171	AHPEFNNSWARIAKGFLPALEYFDRYLHLPRNPQPKYHSQFEPFPSPSNELLD--	228	Qy	423	B-----AVNLEBEKMANLEQD	439		
Qy	210	EAFLMAGKKVATR-FVPGVKAIASSIQDEAK-SESVPKPSRVHATVGETLWKLIAASRAL	267	Db	461	DDRDWVYHHLTKGQLRFEETE	482		
Db	229	-----GKTNKSGTFLFSLRSQEQINTLKQKLDLSNNTTSLSTEVYAAHYWRS-VSKARGL	281	RESULT 13					
Qy	268	TSGTISTRSLIAAQAVNRLTRMMETVLDNATGFLWQAQJLLESLHTPEFSDLK--L	324	US-09-457-046B-63					
Db	282	-SDHEEIKL--IMPVDCRSRINNPSLPKGYCGN-----WFLAVCTATGDSLSCNPL	330	Sequence #3, Application US/09457046B					
Qy	325	CDLYNLINGSVYKQNSDFTETFKGKEGYGRMCYHDFORTMSM-----EPADPIY--	375	Patent No 6287835					
Db	331	TDTAGKVOEALIGLDDYLRS-----AIDHTBSKGLPVPTMGSPPEKTLXPNV	378	GENERAL INFORMATION:					
Qy	376	IPSSWTFN-EPNDEFGNRTSWIGHAGKIESASCKFLILVPTQCGSG	421	APPLICANT: Croteau, Rodney et al.					
Db	379	IVNSWGRIFPYQAMDFGNGSPTFGISNTFYDQGQ--FLIPSRDGDG	422	TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway					
Qy	RESULT 12						FILE REFERENCE: 53679		
Db	US-09-457-046B-65						CURRENT APPLICATION NUMBER: US/09/457,046B		
Qy	Sequence 65, Application US/09457046B						NUMBER OF SEQ ID NOS: 74		
Db	Patent No. 6287835						NUMBER OF SEQ ID NOS: 74		
Qy	GENERAL INFORMATION:						NUMBER OF SEQ ID NOS: 63		
Db	APPLICANT: Croteau, Rodney et al.						LENGTH: 482		
Qy	TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway						TYPE: PRT		
Db	FILE REFERENCE: 53679						ORGANISM: Arabidopsis thaliana		
Qy	CURRENT APPLICATION NUMBER: US/09/457,046B						US-09-457-046B-63		
Db	CURRENT FILING DATE: 1999-12-07						TYPE: PRT		
Qy	NUMBER OF SEQ ID NOS: 74						ORGANISM: Arabidopsis thaliana		
Db	SOFTWARE: PatentIn Ver. 2.1						LENGTH: 497		
Qy	SEQ ID NO 65						TYPE: PRT		
Db	SEQUENCE 65, Application US/09457046B						ORGANISM: Arabidopsis thaliana		
Qy	Query Match 9.4%; Score 222; DB 3; Length 497;						US-09-457-046B-65		
Db	Best Local Similarity 22.9%; Prod. No. 3 Be-15; Mismatches 159; Indels 142; Gaps 28;						Query Match 9.4%; Score 222; DB 3; Length 497;		
Qy	Matches 115; Conservative 86; Mismatches 159; Indels 142; Gaps 28;						Best Local Similarity 22.9%; Prod. No. 3 Be-15; Mismatches 159; Indels 142; Gaps 28;		
Db	55 VITTRKQV-----TRALPQDHWPLSNLDLPLPLNTVFCYKPKLHFTN--TVAYET						Matches 115; Conservative 86; Mismatches 159; Indels 142; Gaps 28;		
Qy	6 VSIINSKHTIKPSTSSTPLQPKVTL--LDQLTTPAYVTFVPTIDHDENFLPQTLA--D						55 VITTRKQV-----TRALPQDHWPLSNLDLPLPLNTVFCYKPKLHFTN--TVAYET		
Db	62 LROALSETLTLYPLSGRYKNN-----LYIDDFERGPYPLLEARVNCUDMTDFLFLRKIEC						6 VSIINSKHTIKPSTSSTPLQPKVTL--LDQLTTPAYVTFVPTIDHDENFLPQTLA--D		
Qy	109 LITPALAETLVSYTAAGFLVTPNTPGEPELNN--RGYDFEVAGDYLRENLYDPPES						62 LROALSETLTLYPLSGRYKNN-----LYIDDFERGPYPLLEARVNCUDMTDFLFLRKIEC		
Db	116 INERFPIKPFSEMAISDERYPLIGQVNVFEDSG-IATGVSYSKHLIDGTADECFLKSGA						109 LITPALAETLVSYTAAGFLVTPNTPGEPELNN--RGYDFEVAGDYLRENLYDPPES		
Qy	167 IAKLVPK-----KHGTAIAOTQQLKCSIVVCTEDERVADAYSMMNELLWSWA						116 INERFPIKPFSEMAISDERYPLIGQVNVFEDSG-IATGVSYSKHLIDGTADECFLKSGA		
Db	175 VPR-----GC-----RENITHPSLARAL---LIPPRDPL--PEKVDQMEAL						167 IAKLVPK-----KHGTAIAOTQQLKCSIVVCTEDERVADAYSMMNELLWSWA		
Qy	217 ISRSDDVPSCVSPRSLLINPRRLVMDPSIDQ1YMPVTSLLPQQTETNPENL--						175 VPR-----GC-----RENITHPSLARAL---LIPPRDPL--PEKVDQMEAL		
Db	213 WFAKKVATRFFVGKRAISSIOEAKSESVPKPSRVHATVGETLWKLIAASRALTSGTT						217 ISRSDDVPSCVSPRSLLINPRRLVMDPSIDQ1YMPVTSLLPQQTETNPENL--		
Qy	270 -----LASRITYTKANALQEL2PLASSRKGRTKQFSFLWKL-LVAEHAKDPVPI						213 WFAKKVATRFFVGKRAISSIOEAKSESVPKPSRVHATVGETLWKLIAASRALTSGTT		
Db	273 ST-RLSIAAQAVNRLTRMMETVLDNATGNLF-----WQAQALIE-----LSHTPRISDLK						270 -----LASRITYTKANALQEL2PLASSRKGRTKQFSFLWKL-LVAEHAKDPVPI		
Qy	323 KTSKLGI-----WVDTGRRRL-MEKENNTYFGNVLSVPGCGRIDDLISKPLSWTVEVH--						273 ST-RLSIAAQAVNRLTRMMETVLDNATGNLF-----WQAQALIE-----LSHTPRISDLK		
Db	324 LCDLUNLINGSVYKQNGDYFETFKGKEGYGRMCYHDFORTMSMSEPAFDI-----TCRPPAVSRIVSYGS						323 KTSKLGI-----WVDTGRRRL-MEKENNTYFGNVLSVPGCGRIDDLISKPLSWTVEVH--		
Qy	376 -----REPKKSVT-----KEHFLNLIDWNE-----TCRPPAVSRIVSYGS						324 LCDLUNLINGSVYKQNGDYFETFKGKEGYGRMCYHDFORTMSMSEPAFDI-----TCRPPAVSRIVSYGS		
Db	RESULT 14						376 -----REPKKSVT-----KEHFLNLIDWNE-----TCRPPAVSRIVSYGS		
Qy	375						376 -----REPKKSVT-----KEHFLNLIDWNE-----TCRPPAVSRIVSYGS		
Db	413 DGAFTVSSGRSPVNOVDFEGNSPFFSSYHFIFWGGSSAG-----YVMPMP---						413 DGAFTVSSGRSPVNOVDFEGNSPFFSSYHFIFWGGSSAG-----YVMPMP---		
Qy	423 B-----AVNLEBEKMANLEQD						423 B-----AVNLEBEKMANLEQD		
Db	461 DDDRWVYHHLTKGQLRFEETE						461 DDDRWVYHHLTKGQLRFEETE		
Qy	462						462		
Db	463						463		
Qy	482						482		
Db	483						483		
Qy	484						484		
Db	485						485		
Qy	486						486		
Db	487						487		
Qy	488						488		
Db	489						489		
Qy	490						490		
Db	491 WGRTSWI-----GVAGKIESASCKFLILVPTQCGSSG-IEAVVNLEBEMANJEQDPHFLAAS						491 WGRTSWI-----GVAGKIESASCKFLILVPTQCGSSG-IEAVVNLEBEMANJEQDPHFLAAS		
Qy	492 WGRPVAVSGRSNKFDGKISAF-----PREGNGTVDLEWVLSETMAGIESDGFEMRVT						492 WGRPVAVSGRSNKFDGKISAF-----PREGNGTVDLEWVLSETMAGIESDGFEMRVT		
Db	493						493		
Qy	494						494		
Db	495						495		
Qy	496						496		
Db	497						497		
Qy	498						498		
Db	499						499		
Qy	500						500		
Db	501						501		
Qy	502						502		
Db	503						503		
Qy	504						504		
Db	505						505		
Qy	506						506		
Db	507						507		
Qy	508						508		
Db	509						509		
Qy	510						510		
Db	511						511		
Qy	512						512		
Db	513						513		
Qy	514						514		
Db	515						515		
Qy	516						516		
Db	517						517		
Qy	518						518		
Db	519						519		
Qy	520						520		
Db	521						521		
Qy	522						522		
Db	523						523		
Qy	524						524		
Db	525						525		
Qy	526						526		
Db	527						527		
Qy	528						528		
Db	529						529		
Qy	530						530		
Db	531						531		
Qy	532						532		
Db	533						533		
Qy	534						534		
Db	535						535		
Qy	536						536		
Db	537						537		
Qy	538						538		
Db	539						539		
Qy	540						540		
Db	541						541		</td

US-09-457-046B-50
 ; Sequence 50, Application US/09457046B
 ; Patent No. 6,287,835
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney et al.
 ; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
 ; FILE REFERENCE: 53679
 ; CURRENT APPLICATION NUMBER: US/09/457,046B
 ; CURRENT FILING DATE: 1999-12-07
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 50
 ; LENGTH: 461
 ; TYPE: PRT
 ; ORGANISM: Taxus cuspidata
 ; US-09-457-046B-50

Query Match 7.5%; Score 179; DB 3; Length 443;
 Best Local Similarity 23.8%; Pred. No. 1.5e-10;
 Matches 101; Conservative 64; Mismatches 176; Indels 84; Gaps 21;

Query 1 MEK----TWSINSKHTIKPSTSSTPLQPKTLTIDQTLTPEP-AXYPIVFFYPTIDHDFN 54
 Db 1 MEKSGSADLWNTNLIBRWWVAF-CQPTBKTLIQLSSTIKNGGFGFANVLLVFG--ASHGVS 56
 Qy 55 LPQTLD---LRQALSETLTLYPLSGRK---NMLYIDDPFEGVPLHARVNCMDMT 106
 Db 57 ----ADPAKTRIREALKTLYVFFYPPAGRLPKKEDGIEVECTIEQALFVMAAND--- 108
 Qy 107 FLRLRKIECLNEFVPPIKPFSNEAIS-DERYV---LIGVONVVF-SGIAIGVSVSHKLID 161
 Db 109 ---LSVYRDLDEYNPLERQLOSSLSDTDXDHLMTVQVTPETCGFVNGTSVHOSICD 165
 Qy 162 GGTALDFLKSWGAVPFGCRENITHPSLSEALLPFPRDDLP-EKTYDQMLWFAKGKVA 220
 Db 166 GNGLGFFKSNAEIVGEVKSIEPWNRE--LvkPbDyHfQJLQVSEFIRPLVVEKVG 223
 Qy 221 TRRFVPGVKAQSLIQDEAKSSSVPKPSRVRHAYTGFELWKLHLASRALTSGTSTRSIAA 280
 Db 224 QTSLVTSFEKINHICRCEMEKESESSEFVTAATMWW---LARTRAFOQIPEN----- 272
 Qy 281 QAVNTRMNNETVLD---NATGNIFWNAQAILBLISHTYPEISDILKLCJLNLINGS- 334
 Db 273 EDVTLIAMDARRSFDEPIPKGGYGTW-----GTYAKD---NvHNLJSGSL 317
 Qy 335 -----VKOCNGDYFETFGKIGEGYGERMCYXLDFOFTTMSMPAPDYLFSWTNF-FNPL 387
 Db 318 DHALTVTKNSMSSFYENMTSSR-----VLYNPSTLDSMKYENTVLSLDSRLGHNEV 369
 Qy 388 DEFGWG 392
 Db 370 DEFGWG 374

RESULT 15
 US-09-457-046B-70
 ; Sequence 70, Application US/09457046B
 ; Patent No. 6,287,835
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney et al.
 ; TITLE OF INVENTION: Transacyclases of the Paclitaxel Biosynthetic Pathway
 ; FILE REFERENCE: 53679
 ; CURRENT APPLICATION NUMBER: US/09/457,046B
 ; CURRENT FILING DATE: 1999-12-07
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 70
 ; LENGTH: 461
 ; TYPE: PRT
 ; ORGANISM: Arribidopsis thaliana
 ; US-09-457-046B-70

Query Match 7.5%; Score 178.5; DB 3; Length 461;

Best Local Similarity 24.3%; Pred. No. 1.8e-10;
 Matches 118; Conservative 68; Mismatches 197; Indels 103; Gaps 26;
 Qy 1 MEKISVINSKHTIKPSTSSTPLQPKTLTIDQTLTPEP-AXYPIVFFYPTIDHDFNQTLA 60
 Db 13 LERKPVBL---VPS-KTHCCTLSLSDN-DPENEYMYATIVYFKANGKNDPPVS 65
 Qy 61 DLQMLSETLTLYPLSGRK---KNNYTIDDFBEGVPLHARVNCMDTFLRLRKIEC 115
 Db 66 LIRKALSELVAVHYPLSGKLMRSBZERGKLVVLSGVPFVATSLDSS--LYNIEN 122
 Qy 116 LNEFVPIK--PFSMEAIISDERYPLQVQVNF-SGIAIGVSVSHKLID-GTADCFLK 170
 Qy 123 LDQVQALRLVPEIDYESNWCYHPLAQVTKFACGGFTGTLATHAVCDGKGYA----- 177
 Db 171 SWGAVERGCRENITHPSLSEAL-----1FPPRDLPEK---YVQMEA 211
 Db 178 -----QIIH-ALTELAAGK7EPSVKSIVQBRVLVGKIDNPKVPGSH1DGFIA 225
 Qy 212 --LWFENGKVKATRREVFGVRAISSID-----EAXESVYKPSRHYATVPLKHLIAA 263
 Db 226 TSAYLTEDVVTETINRAGD1KRLKD3MMKECEYLIKESF---TIVYELSSYIWK--LR 279
 Qy 264 SRALTSGTSTRSLIAAQAVNLRTRNNMETVLD-----NATGNIFWNAQAILELSHTPE 318
 Db 280 SRALKUNDPG--ITVGLVGNSTR-----HYLDPLPLPGYGNAY-1DVYVFL--TVRE 327

Db 319 ISDLKUDCDVNLNGSYKQC--NGDYFETFGKIGEGYGERMCYXLDFOFTTMSMKEAPD-IV 375
 Db 328 LBESSIONTIANEWKAKKTAVBKGYIEELKNT-----RIMRDDSMPFEGVSPGLF 378
 Qy 376 LFSSWTN--FFNPLDFGWRGTRSWIGVAGKIBASCRFILVPTQCGGIEAWVN--LEEE 431
 Db 379 FLIDWANIGWFGSMDFQWNEPWNRLFTORESTVWGMILKPSKSDSMEGGVXVTKLP 438
 Qy 432 KMMAL 437
 Db 439 RDAMVE 444

Search completed: July 6, 2004, 13:40:15
 Job time : 20.1722 secs

Tue Jul 6 14:13:41 2004

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 13:38:51 : Search time 49.964 Seconds

(without alignments)
2816.033 Million cell updates/sec

Title: US-09-857-518A-6

Perfect score: 1276540

Sequence: 1 MEKIEVSINSKHTIKPSTSS.....MAMLEQDPHFLLASPKTLI 452

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 311283816 residues

Total number of hits satisfying chosen parameters: 1276540

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Published Applications AA: *

```

1: /cgn2_6/_ptodata/2/_pubpaal/us07_PUBCCOMB_.pep:*
2: /cgn2_6/_ptodata/2/_pubpaal/PCT_NEW_PUB_.pep:*
3: /cgn2_6/_ptodata/2/_pubpaal/us05_NEW_PUB_.pep:*
4: /cgn2_6/_ptodata/2/_pubpaal/us06_PUBCCOMB_.pep:*
5: /cgn2_6/_ptodata/2/_pubpaal/us07_NEW_PUB_.pep:*
6: /cgn2_6/_ptodata/2/_pubpaal/ECTOS_PUBCCOMB_.pep:*
7: /cgn2_6/_ptodata/2/_pubpaal/us08_NEW_PUB_.pep:*
8: /cgn2_6/_ptodata/2/_pubpaal/us08_PUBCCOMB_.pep:*
9: /cgn2_6/_ptodata/2/_pubpaal/us09_PUBCCOMB_.pep:*
10: /cgn2_6/_ptodata/2/_pubpaal/us09B_PUBCCOMB_.pep:*
11: /cgn2_6/_ptodata/2/_pubpaal/us09C_PUBCCOMB_.pep:*
12: /cgn2_6/_ptodata/2/_pubpaal/us10A_PUBCCOMB_.pep:*
13: /cgn2_6/_ptodata/2/_pubpaal/us10B_PUBCCOMB_.pep:*
14: /cgn2_6/_ptodata/2/_pubpaal/us10C_PUBCCOMB_.pep:*
15: /cgn2_6/_ptodata/2/_pubpaal/us10C_PUBCCOMB_.pep:*
16: /cgn2_6/_ptodata/2/_pubpaal/us10C_PUBCCOMB_.pep:*
17: /cgn2_6/_ptodata/2/_pubpaal/us60_NEW_PUB_.pep:*
18: /cgn2_6/_ptodata/2/_pubpaal/us60_PUBCCOMB_.pep:*

```

Database :

1 US-09-866-5172A-66
 2 US-09-866-5172A-66
 3 US-09-866-5170A-62
 4 US-09-866-5170A-62
 5 US-09-866-5170A-62
 6 US-09-866-5170A-62
 7 US-09-866-5170A-62
 8 US-09-866-5170A-62
 9 US-09-866-5170A-62
 10 US-09-866-5170A-62
 11 US-09-866-5170A-62
 12 US-09-866-5170A-62
 13 US-09-866-5170A-62
 14 US-09-866-5170A-62
 15 US-09-866-5170A-62
 16 US-09-866-5170A-62
 17 US-09-866-5170A-62
 18 US-09-866-5170A-62
 19 US-09-866-5170A-62
 20 US-09-866-5170A-62
 21 US-09-866-5170A-62
 22 US-09-866-5170A-62
 23 US-09-866-5170A-62
 24 US-09-866-5170A-62
 25 US-09-866-5170A-62
 26 US-09-866-5170A-62
 27 US-09-866-5170A-62
 28 US-09-866-5170A-62
 29 US-09-866-5170A-62
 30 US-09-866-5170A-62
 31 US-09-866-5170A-62
 32 US-09-866-5170A-62
 33 US-09-866-5170A-62
 34 US-09-866-5170A-62
 35 US-09-866-5170A-62
 36 US-09-866-5170A-62
 37 US-09-866-5170A-62
 38 US-09-866-5170A-62
 39 US-09-866-5170A-62
 40 US-09-866-5170A-62
 41 US-09-866-5170A-62
 42 US-09-866-5170A-62
 43 US-09-866-5170A-62
 44 US-09-866-5170A-62
 45 US-09-866-5170A-62

Sequence 66, Appl

RESULT 1
 US-09-866-5172A-62
 ; Sequence 62, Application US/098665172A
 ; Patent No. US2002138859A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Croteau, Rodney et al.
 ; TITLE OF INVENTION: Transcytases of the Paclitaxel Biosynthetic Pathway
 ; FILE REFERENCE: 53679
 ; CURRENT APPLICATION NUMBER: US/09/866,572A
 ; PRIORITY APPLICATION NUMBER: 09/457,046
 ; PRIORITY FILING DATE: 1999-12-07
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 62
 ; LENGTH: 436
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana

US-09-866-5172A-62
 Query Match 21.0%; Score 498; DB 9; Length 436;
 Best Local Similarity 30.9%; Pred. No. 8.9e-44; Indels 184; Mismatches 184; Gaps 19;
 Matches 145; Conservative 90; Nucleotides 184; Indels 50; Gaps 19;

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	498	21.0	436	9 US-09-866-5172A-62	Sequence 62, Appl
2	498	21.0	436	9 US-09-866-5170A-62	Sequence 62, Appl
3	498	21.0	436	14 US-10-166-984-62	Sequence 62, Appl
4	498	21.0	436	15 US-10-165-984-62	Sequence 62, Appl
5	485	20.4	446	9 US-09-866-5172A-74	Sequence 74, Appl
6	485	20.4	446	9 US-10-166-984-74	Sequence 74, Appl
7	485	20.4	446	14 US-10-166-984-74	Sequence 74, Appl
8	485	20.4	446	14 US-10-166-984-74	Sequence 74, Appl
9	448	18.9	435	9 US-09-866-5170A-60	Sequence 60, Appl
10	448	18.9	435	9 US-10-166-984-60	Sequence 60, Appl
11	448	18.9	435	14 US-10-166-984-75	Sequence 75, Appl
12	448	18.9	435	15 US-10-166-984-75	Sequence 75, Appl
13	438	18.5	474	12 US-10-167-603C-7	Sequence 7, Appl
14	438	18.5	474	12 US-10-167-603C-14	Sequence 14, Appl
15	438	18.5	474	12 US-10-167-603C-16	Sequence 16, Appl

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution. and is derived by analysis of the total score distribution.

SUMMARIES

Query Match	Score	498; DB 9; Length 436;
Qy	21.0%	Score 498; DB 9; Length 436;
Qy	30.9%	Best Local Similarity 30.9%; Pred. No. 8.9e-44; Indels 184; Mismatches 184; Gaps 19;
Qy	90	Matches 145; Conservative 90; Nucleotides 184; Indels 50; Gaps 19;
Qy	1	MEKTVINSRKRTHIKPSTSSTP--LQPYKLTLLDOLTPRXPVYFEPYPTDHD--FNL 55
Qy	1	MEKTVINSRKRTHIKPSTSSTP--LQPYKLTLLDOLTPRXPVYFEPYPTDHD--FNL 55
Qy	1	MEK-TWELLREIVRS-SPTPDKRILNLSLDDLSSPMTGALFYAADPQNLGFST 58
Db	1	PQTLADLROAISSETLTYPLSGRYVNNLYTDDFEGGPVYIPEARNCDMDFLRLRKEC 115
Qy	56	ETSLRILKSKSISKTLFYPAGRIGS-FVCENDGAEVFEARTVLLFELISESEFLKCYPES 117
Db	59	ETSLRILKSKSISKTLFYPAGRIGS-FVCENDGAEVFEARTVLLFELISESEFLKCYPES 117
Qy	116	LNEFPIKPPFSEALIFPDKLTDGKTVTRREYFGVXKAISSI 234
Qy	118	LELLTVEAKCREAVT--WPDLLIIDANFEESGELVITCVSHKCTDATSLAMPINGWAE 174
Qy	175	VFRGCRENTIIFPSLSEALIFPDKLTDGKTVTRREYFGVXKAISSI 234

334 TNAVNTINGTSSSESSQVILHEDFKDKGSLEKVIQNFASGHD-----388

363 RTMSSKEPAPDI---YLFSSWTFNPFNPLDFGWTSGTSMWIGVAGKIE-SASCKPFLVPT 416

389 ---ASIKKINDVEVINFIWISSWCRMGLYEIDEGWKPITVDENIKPNKNCFF--MNDT 443

417 QCGSGIEAWVNLEEKRAMLE 437

444 KCGEGIEWASFEDMAKFE 464

RESULT 15

US-10-167-603C-16

Sequence 16, Application US/10167603C

Publication No. US20030226161A1

GENERAL INFORMATION:

APPLICANT: ZENK, Meinhardt H.

APPLICANT: KUTCHAN, Toni

APPLICANT: GROTH, Torsten

TITLE OF INVENTION: Salutaridinol 7-O-Acetyltransferase and derivatives

FILE REFERENCE: 67529

CURRENT APPLICATION NUMBER: US/10/167,603C

CURRENT FILING DATE: 2002-06-11

PRIOR APPLICATION NUMBER: EP 01114142.3

PRIOR FILING DATE: 2001-06-11

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 16

LENGTH: 474

TYPE: PRT

ORGANISM: Papaver somniferum

US-10-167-603C-16

Query Match 18.5% Score 438; DB 12; Length: 474;

Best Local Similarity 27.9%; Pred. No. 2, 6e-37; Indels 104; Gaps 21

Matches 140; Conservative 87; Mismatches 170; Insertions 104; Gaps 21

Qy 1 MEKIEWSINSKHTKIPST-SSTPLQPKLTLLDOLTP-PAYVPIVFFYPT-----49

Db 4 MYSAAEVISKEKIPPTETPSQNLNFNLSSLDQCPFLYYVPLFYPTAANSGTSSN 63

Qy 50 DHDENLPQTLADLRLQALSETLTYYPLSGRVRLNLYLDFEGVPLYLERVNCDMTDPLR 109

Db 64 HED---DLDLKLSSSLKRTVTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 257

Db 64 HAD---DLDLKLSSSLKRTVTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 257

Db 64 HAD---DLDLKLSSSLKRTVTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 257

Db 110 IRKIECLNEFVPIKPFMSMEAISDERYPLIGQVNVDF-SGIAIGVSVSHKLIDGTADCF 168

Db 119 QDPVP-LSQLPSEVVS--ASPKEALVQVNMEDGCTAACSSVSKHIAADATMSTF 174

Db 169 LKSGMELWFGAKKVRTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 205

Db 175 IRSWASTTKTSRGGSTAAVTQDQLIPSP-DSASLFLPPSERLTSPSGMSEIPFSSTPDT 233

Db 234 EDD-----KTVSKRFVDFAKITSYREKLQVLMHDNYKSR--RQTRVEVTSLLN 281

Db 258 XHLIAASRALSGTTSRSLIAAOVNLTRAMMETVLDNATGFLFWWAQALELSHTT-316

Db 282 K-----SVMKSTPAGLIPVVAHVNLRKMD-PLQDVSFGNLSVTSALPATTIT 333

Db 317 -----PEISDLKLCDLVLNLLSVPQCNQDXYFETPKGEGYGRMCYLDFO 362

Db 334 TNAVNTINGTSSSESSQVILHEDFKDKGSLEKVIQNFASGHD-----388

Qy 363 RTMSSKEPAPDI---YLFSSWTFNPFNPLDFGWTSGTSMWIGVAGKIE-SASCKPFLVPT 416

Qy 389 ---ASIKKINDVEVINFIWISSWCRMGLYEIDEGWKPITVDENIKPNKNCFF--MNDT 443

Qy 417 QCGSGIEAWVNLEEKRAMLE 437

Db 444 KCGEGIEWASFEDMAKFE 464

RESULT 16

US-10-167-603C-16

Sequence 16, Application US/10167603C

Publication No. US20030226161A1

GENERAL INFORMATION:

APPLICANT: KUTCHAN, Toni

APPLICANT: GROTH, Torsten

TITLE OF INVENTION: Salutaridinol 7-O-Acetyltransferase and derivatives

FILE REFERENCE: 67529

CURRENT APPLICATION NUMBER: US/10/167,603C

CURRENT FILING DATE: 2002-06-11

PRIOR APPLICATION NUMBER: EP 01114142.3

PRIOR FILING DATE: 2001-06-11

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 16

LENGTH: 474

TYPE: PRT

ORGANISM: Papaver somniferum

US-10-167-603C-16

Query Match 18.5% Score 438; DB 12; Length: 474;

Best Local Similarity 27.9%; Pred. No. 2, 6e-37; Indels 104; Gaps 21

Matches 140; Conservative 87; Mismatches 170; Insertions 104; Gaps 21

Qy 1 MEKIEWSINSKHTKIPST-SSTPLQPKLTLLDOLTP-PAYVPIVFFYPT-----49

Db 4 MYSAAEVISKEKIPPTETPSQNLNFNLSSLDQCPFLYYVPLFYPTAANSGTSSN 63

Qy 50 DHDENLPQTLADLRLQALSETLTYYPLSGRVRLNLYLDFEGVPLYLERVNCDMTDPLR 109

Db 64 HED---DLDLKLSSSLKRTVTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 257

Db 64 HAD---DLDLKLSSSLKRTVTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 257

Db 64 HAD---DLDLKLSSSLKRTVTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 257

Db 110 IRKIECLNEFVPIKPFMSMEAISDERYPLIGQVNVDF-SGIAIGVSVSHKLIDGTADCF 168

Db 119 QDPVP-LSQLPSEVVS--ASPKEALVQVNMEDGCTAACSSVSKHIAADATMSTF 174

Db 169 LKSGMELWFGAKKVRTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 205

Db 175 IRSWASTTKTSRGGSTAAVTQDQLIPSP-DSASLFLPPSERLTSPSGMSEIPFSSTPDT 233

Db 234 EDD-----KTVSKRFVDFAKITSYREKLQVLMHDNYKSR--RQTRVEVTSLLN 281

Db 258 XHLIAASRALSGTTSRSLIAAOVNLTRAMMETVLDNATGFLFWWAQALELSHTT-316

Db 282 K-----SVMKSTPAGLIPVVAHVNLRKMD-PLQDVSFGNLSVTSALPATTIT 333

Db 317 -----PEISDLKLCDLVLNLLSVPQCNQDXYFETPKGEGYGRMCYLDFO 362

Db 334 TNAVNTINGTSSSESSQVILHEDFKDKGSLEKVIQNFASGHD-----388

Qy 363 RTMSSKEPAPDI---YLFSSWTFNPFNPLDFGWTSGTSMWIGVAGKIE-SASCKPFLVPT 416

Qy 389 ---ASIKKINDVEVINFIWISSWCRMGLYEIDEGWKPITVDENIKPNKNCFF--MNDT 443

Qy 417 QCGSGIEAWVNLEEKRAMLE 437

Db 444 KCGEGIEWASFEDMAKFE 464

RESULT 17

US-10-167-603C-16

Sequence 17, Application US/10167603C

Publication No. US20030226161A1

GENERAL INFORMATION:

APPLICANT: KUTCHAN, Toni

APPLICANT: GROTH, Torsten

TITLE OF INVENTION: Salutaridinol 7-O-Acetyltransferase and derivatives

FILE REFERENCE: 67529

CURRENT APPLICATION NUMBER: US/10/167,603C

CURRENT FILING DATE: 2002-06-11

PRIOR APPLICATION NUMBER: EP 01114142.3

PRIOR FILING DATE: 2001-06-11

NUMBER OF SEQ ID NOS: 19

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 16

LENGTH: 474

TYPE: PRT

ORGANISM: Papaver somniferum

US-10-167-603C-16

Query Match 18.5% Score 438; DB 12; Length: 474;

Best Local Similarity 27.9%; Pred. No. 2, 6e-37; Indels 104; Gaps 21

Matches 140; Conservative 87; Mismatches 170; Insertions 104; Gaps 21

Qy 1 MEKIEWSINSKHTKIPST-SSTPLQPKLTLLDOLTP-PAYVPIVFFYPT-----49

Db 4 MYSAAEVISKEKIPPTETPSQNLNFNLSSLDQCPFLYYVPLFYPTAANSGTSSN 63

Qy 50 DHDENLPQTLADLRLQALSETLTYYPLSGRVRLNLYLDFEGVPLYLERVNCDMTDPLR 109

Db 64 HED---DLDLKLSSSLKRTVTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 257

Db 64 HAD---DLDLKLSSSLKRTVTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 257

Db 64 HAD---DLDLKLSSSLKRTVTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 257

Db 110 IRKIECLNEFVPIKPFMSMEAISDERYPLIGQVNVDF-SGIAIGVSVSHKLIDGTADCF 168

Db 119 QDPVP-LSQLPSEVVS--ASPKEALVQVNMEDGCTAACSSVSKHIAADATMSTF 174

Db 169 LKSGMELWFGAKKVRTRFVGKTAISIQ-----DEAKSESPVKPSRVHATGFLW 205

Db 175 IRSWASTTKTSRGGSTAAVTQDQLIPSP-DSASLFLPPSERLTSPSGMSEIPFSSTPDT 233

Db 234 EDD-----KTVSKRFVDFAKITSYREKLQVLMHDNYKSR--RQTRVEVTSLLN 281

Db 258 XHLIAASRALSGTTSRSLIAAOVNLTRAMMETVLDNATGFLFWWAQALELSHTT-316

Db 282 K-----SVMKSTPAGLIPVVAHVNLRKMD-PLQDVSFGNLSVTSALPATTIT 333

Db 317 -----PEISDLKLCDLVLNLLSVPQCNQDXYFETPKGEGYGRMCYLDFO 362

Db 334 TNAVNTINGTSSSESSQVILHEDFKDKGSLEKVIQNFASGHD-----388

Qy 363 RTMSSKEPAPDI---YLFSSWTFNPFNPLDFGWTSGTSMWIGVAGKIE-SASCKPFLVPT 416

Qy 389 ---ASIKKINDVEVINFIWISSWCRMGLYEIDEGWKPITVDENIKPNKNCFF--MNDT 443

Qy 417 QCGSGIEAWVNLEEKRAMLE 437

Db 444 KCGEGIEWASFEDMAKFE 464

Search completed: July 6, 2004, 13:48:41
Job time : 50.964 secs

RESULT 3	
Db	EF11418
	hypothetical protein - Arabidopsis thaliana
C:Species:	Arabidopsis thaliana (mouse-ear cress)
C:Variety:	Colombia
C:Accession:	EF11418
R:Bevan, M.; Bancroft, I.; Bent, B.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dijk, P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitznerger, T.; Pohl, T.M.; Terryn, N.; Giel, avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.	
Nature 391, 485-488, 1998	
A:Authors:	Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pors, A.; Puttemans, P.; Ansechtoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Paine, K.; Benes, V.; Rechman, S.; Ans
A:Title:	Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number:	AT1400; MUID:9812113; PMID:9461215
A:Accession:	EF11418
A:Status:	preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type:	DNA
A:Residues:	1-435 <BEV>
A:Cross-references:	GB:Z97338; NID:92244870; PID:e326526; PID:92244897
A:Genetics:	
A:Map position:	4COP9-4G3845
Query Match	18.9%
Best Local Similarity	29.8%
Matches	137; Conservative 79; Mismatches 193; Indels 50; Gaps 15;
Qy	3 KIEVNSINSKHTIKPSSTSPLQPKWLTLDQLTTPAYVPIVFFYPTTDHDENLPCLTIALD 62
Db	4 KLEVTT--GKGVKPA-SESPRDRQGKLSLILYCPQIYVSTIFFYLITESSEV-FSENL 58
Qy	63 RQALSETTLYPLSPGRVKNNLYIDFEQGPYPLEARVNCDMTDFRLRKCECLNBFVPI 122
Db	59 KLSLSSETLSRSPYPLGRIE GEJSICNDEGAVFTARTDILLPDRFLNLTDSLSGFLPT 117
Qy	123 KPFMSMEASIDERYPLPGYVNNF- DSGTAGVSVYSHKLIDGTTACDFLKSWGAVFRGCR 180
Db	118 ---LAGESSPAAPWPLSVKTVFFGGSGVAVSVSHKCIDIASLVTFVDWATTAAGK 174
Qy	181 EN-----IHPPLSEAAALLPPRDR--DLPKRYVQDQMEALNPAGKKVATREFVFGV 228
Db	175 SNSTIEAETIYPPPESEMYEQPSTSDSNITKTV-----LKRFPVPEP 220
Qy	229 KAISSJOADEKSEVSPKPSRVHAYVFLWXLHIANRALSUSGTTSPRLSIAQAVMNLTR 288
Db	221 SKIAELKHKAASESPVPIRVEAISLWR--CRNNSSSSNLIRQAVWQAMETRLR 277
Qy	269 MNMETFLDNATGNIFLWWAQALIELSHTTPEISDLGLCLVNLNGSSVKCNGDYFTFKG 348
Db	278 IPSSVAPDQIGNL---SFGFLSKDAS--SEFEPEIATFRKNKERNEMIKESLQG 331
Qy	349 KEGYGRMCYELDFQRTMSKSEKAPAPDYLESSWNTF- FNPLDFGNGRTSWIGVAGKIESAS 407
Db	332 NT-1CQSLSLSLMAETVSESTE--IIRYINSSWCRKPFYEVDFGSSSPVWGYASHTYDN 388
Qy	408 CKFTLILVPTCGSGI-BAWNLYEEBNAMEQDPHFLALA 446
Db	389 MGVVILDSKREGDGYTEAMWISLPEEMSIVYDDQELAYA 427

A; Reference number: Z23008	Matches 114; Conservative 88; Mismatches 185; Indels 102; Gaps 26;
A; Accession: T45216	Qy 4 IEVSINSKHTIKPSTSTPLQPKYKUTLJUDQL--TPPAVYPIVFFYPTTDH-----51
A; Status: Preliminary	Db 1 MSIQIKQSTMTVPR-AEPEPNKSILWLSKIDMTRPTYSITGAVIYKOPPNNDNIHESS 59
A; Molecule type: DNA	Qy 52 ---DFNLQTLADLQLQASSETLTLYPLSPGRVAN--LYIDDEEGVPLLEARVCDMT 105
A; Residues: 1-430 <CHO>	Db 60 MYDANI -----LIEALSKAIVPPYPPAGRKINGDREYIDCNAEGALFVEAESHYLIE 113
A; Cross-references: EMBL:AU133315	Qy 106 DEFLRLKIECLINE-FVPIKPSFNEAISDERYPLGVQVNVD-SGIAVCFVSHKLIDIG 163
A; Experimental source: cultivar Columbia; BAC clone T8P19	Db 114 DEGDFPFDDELHRYMPVCDYS-KGIGS -PPLMMQLTRFRCCGVSIGAQMHHACDMG 170
C; Genetics:	Qy 164 TADCFLKSWGAVAFRGCRENI--IHPSLSEAALLFPPR-----DDLPKPVYDPM 209
A; Map position: 3	Db 171 SHPEFNNSWARIANGLILDAEPVHDYRJHLRLNPPQIKYTHSQFEPFVPSLMPNLD - 228
A; Introns: 144/3	Qy 210 EALWMPAGKVKATR -PVFGVKALSSIIQD -EAKSERSPKPSRYHATVGTFLWKLIAASRA 266
A; Note: T8P19.230	Db 229 -----GKTNKSKTOLFLKLSRQONTLKGKLDISNTTRLSYEVAGHWRs-VSKARG 281
C; Superfamily: Arabidopsis CER2 protein	Db 267 LTSGTTPLSLIAQAAVNLTRMMTEVLDNAAGNLFWQAALLESHTTPEISDLKLCD 326
Query Match 11.0%; Score 262; DB 2; Length 430;	Db 282 L-SDHEIKL---IMPVDGSRINNSPLPKGYCGN-----WVFLAVTATVGDLS-CN 329
Best Local Similarity 24.5%; Pred. No. 5.7e-14;	Qy 211 FMSNEAISDERYPLGVQVNVD-SGIAVCFVSHKLIDGTAFCPLKSWGAVFRGCRENI 183
Matches 115; Conservative 72; Mismatches 176; Indels 106; Gaps 23;	Db 229 LYVNLNGSVYKOCNGDYFETFKGKEGYGMCEYLDFORMSMSEPAFDI-----374
A; Introns: 144/3	Qy 270 YLFFSNTNF -FNPFLDFGKGRTSWIGVAGKLESASCKFILVPTQCGSG-IEAWVNLE 429
Db 18 PPVSETPQHYYLNSLNDQ-NIATIVKLYYYKSBSRTNQESTXN-----IKRKLSELV 70	Db 310 PLDTAGKQ-----EALKGQD-----DYL -RSLAIDHTSKSPDLPVPMGSPKETL 375
Qy 16 PSTSSTPLQPKYKLTLDQTLPPAYVPIVFFYYP ---ITDHDPLNLPOTLADLQLQASSTLT 71	Qy 375 --YLPNLVNSGRITYQAMDFGNGSPTFGGSNTFYDQGQ---FLIPSONQDGSMTLAINLF 432
Db 17 KXYPLSLSRV---KNNLYIDDFEGVPLKEAFTNC---DMTFDIFRLKIECLNEFVPIKP 124	Db 376 YPNVLVNSGRITYQAMDFGNGSPTFGGSNTFYDQGQ---FLIPSONQDGSMTLAINLF 432
Qy 17 HYPVAGRLTISEGKFLAVNCTGEGVVVAEANCGIDTIASENTRMTEKLAYTDV 130	Qy 430 EEXNAMLEQ 438
Db 184 THFSLRRAALI---FPPRDLPLPEKVYDOMEALWFAKG---KVATRFRFVGKAISIQ 235	Db 433 SSHUSLFLKK 441
Qy 188 131 GARNIL---EPIPQVYVQTNFKCGFVFLGLGMNSHNTGVAEEFLNSWCENAKGLPLSV 187	RESULT 6
Db 188 - -PPFLDRTILRSRNPKPLTEPNEFOZIEDSTGKLYDEEKLTKSFLFPEPEKLEKLK 245	Db 10918 hypothetical protein F21B7.32 - Arabidopsis thaliana
Qy 236 DEAKSE-SVPKRSRVAHTGFLWKLTAASRAITSGTISTRSLIAQAAVNLTRM-----289	C; Species: Arabidopsis thaliana (mouse-ear cress);
Db 246 TMAIEENNKKVYSTFOALTFGLWLRKSRCAALR - FKPDRQVLFKLLFAADG --RSRFIFPLP 300	C; Date: 12-Feb-1999 #sequence_revision 12-Feb-1999
Qy 290 -----NMETVLDNATGNGLFWWAQAALELSHTTPEISDLKLCDLVNLNGSVKQC 338	C; Accession: T099.18
Db 301 OGKCGNGIVLGLVTSSEBLVGN-----PLISHTYGLVK----RLVLEYTDGFMRs 346	R; Shihinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Coe, J.R.; Eckert, A.; Ecker, J.R.
Qy 339 NEDYFETFKGKEGYGRMCBEYLDFQRTMSSMPEAPDIYLFSSWTFN-FNPLDFGNGRTSW- 396	submitted to the EMBL Data Library, January 1998
Db 347 AMDYFEV-----NRTRPSMNT--LLITMSKLTLEKLDLFGNGEPVFS 387	A; Description: Genomic sequence for Arabidopsis thaliana BAC F21B7.
Qy 397 - -TGVACKTIESACKFLVPTQCGSG---TEAWVNLEESQVAMLEQ 438	A; Reference number: Z14208
Db 388 GPVGLPGR-----EVTFDLP-----SGDDMKSINVFLGPTSAMEVPEE 426	A; Accession: T099.18
RESULT 5	A; Status: Translated from GB/EMBL/DBJ
T10711 anthranilate N-benzoyltransferase (EC 2.3.1.144) - clove pink	A; Molecule type: DNA
N; Alternative names: anthranilate N-hydroxycinnamoyl/benzoyltransferase	A; Residues: 1-461 <SH1>
C; Species: Dianthus caryophyllus (clove pink)	A; Cross-references: EMBL:AC002560; NID:92618677; PID:92809263; GSPDB:GN00059; ATSP:F21B
C; Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000	C; Genetics:
C; Accession: T10711	A; Map position: 1
A; Status: Translated from GB/EMBL/DBJ	Qy 4 IEVSINSKHTIKPSTSTPLQ-----YKLTLFIDQLTPAYVFFYPTIDDFNLFQ 57
R; Yang, Q.; Mattern, U.; Grimmig, B.	Db 15 IVPVINQQFLVHPRs-SPPTANQSPHSLYLSNLIDIGARVFTPSVYFPTSNRRESF-- 71
Plant Mol. Biol. 38, 1201-1214, 1998	Qy 58 TIAIDRQALQASBTLTLYPLSPGRVK -NNLYIDDF --EGVPLLEARVCDMTDLRLK 112
A; Title: Anthranilate N-hydroxycinnamoyl/benzoyltransferase gene from carnation: Rapid	Db 72 VLRKLQDALSEVLYPLSPGRVKREVNENSLKEVFGEEQSVLMSANSNSMDLAD-----125
A; Reference number: Z17092; PMID:9084770; PMID:93288180	Qy 113 IECINNEFVPIKPSMEAI---SDERY----PLIGVQVANED-SGIAIGVSVSHKLIDG 162
Q; Keywords: catalyzes the synthesis of anthranilate	Best Local Similarity 10.2%; Score 242; DB 2; Length 446;
C; Pathway: phytalexin biosynthesis	Best Local Similarity 23.3%; Pred. No. 2.9e-12;

126	--LGDLTVPRPAMPLIFLFRPGEBAYKILMPLIAQQTFTGGSFLGLRLCHCICDCG	182	Db	370	--VSDVLFIAF-----IDFGIWGKPVYGGIGTAGVEDLPGASFYVSFERNNEIGIV	419
163	GTADCFIKSKGAVFRGCRENTIHP-----LSEAAALLFPPRDDLPEKTYDQMEA	211	Qy	424	AWTNEEKKM	433
183	FGAMQFLGGSAAATAK-TGKLTADPEPVWDRTEFKPENPMPYKPHHEYLPIEERSNNTS	241	Db	420	VPTVCLPERCAM	429
212	LWFAKGKV-----AFTERFVGKVAIISIQQDAEAKSESVPKPSRV----HATVGFLLGHLIA	262	Qy	212	RESULT 8	
242	LWDT-KPLQKCYRISKFQCRVKSILQED-----PTLVCSTPDAAMAHIMFWSW-	290	Db	212	anthranilate N-benzoyltransferase (EC 2.3.1.144) (clone pchbct3) - clove pink	
263	ASRALTSQTTSSTRLSAAOAVNLRTTMMMBTYLDNATGFLNFMWQAATL-----ELSH	314	Qy	263	N:Alternative names: anthranilate N-hydroxycinnamoyl/benzoyltransferase	
291	-KALDVKPLDYNLRTF-SVNRVTRTLETLKRGPGYGVCLACMSYESLINDLSK	347	Db	291	C:Species: Dianthus Caryophyllus (clove pink)	
315	TTPESDLKLCDLVNLNLNGSYKQNCDFEFPKGKRGYGRMCYELDFQRTMSSMEAPDI	374	Qy	315	C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000	
348	348 TTTRLYQDARL-----RVSBDYLR-----MVDYDVVKRP-KRLEFGKRL	385	Db	348	C:Accession: T10719	
375	YLFSSTNTNF--FNPDLDEGRGRTSNGVAGKLE-SASACKFTLIVPTOCGSGTEA	424	Qy	375	C:Title: Characterization and heterologous expression of hydroxycinnamoyl/benzoyl-CoA:O-lyase L.	
386	TI-TQWTREMYETADFGWGPVY--AGPIDLRLPFPQCVLILPQ--GGVYES	431	Db	386	A:Reference number: 217095;	
			Qy	386	A:Accession: T10719	
			Db	386	A:Status: preliminary; translated from GB/EMBL/DDBJ	
			Qy	386	A:Molecule type: mRNA	
			Db	386	A:Residues: 1-445 <XAM>	
			Qy	386	A:Cross References: EMBL:284571; NID:92239090; PMID:92239091; PID:92239091	
			Db	386	C:Function:	
			Qy	386	A:Description: catalyzes the synthesis of anthranilate	
			Db	386	A:Pathway: phytocalixin biosynthesis	
			Qy	386	C:Keywords: acyltransferase; coenzyme A	
			Qy	386	Query Match 9.9%; Score 234.5; DB 2; Length 461;	
			Db	386	Best Local Similarity 23.5%; Pred. No. 1.3e-11; Mismatches 189; Gaps 27;	
			Qy	386	Matches 114; Conservative 92; Mismatches 189; Indels 91;	
			Db	386	Qy 4 IEVSINSRHTIKEPKSTSPLQPKYLIDQL--TPPAPYPIVFFYPTIDHDENLPQTLLAD 61	
			Db	386	Db 1 MSHIKKSTGMVRA-EETPNKSDMWSKIDMURYKQPKLNNEIDQPKSS 59	
			Qy	386	Qy 62 -----LROA1SETLTLYXPLISGRGVKNN--LYTDDPFEGYDYLEARNMCDMDFRLR 111	
			Db	386	Db 60 MYFDANILK11AL1KALVYPPMAGRLKINGDRYEDCNGEGALVYEAESHVLDFGDFR 119	
			Qy	386	Qy 112 KIECLNL-FVPLKPFKSMRASDERYPLLGQVQNFID-SCGAIATSVSHALIDGTACFL 169	
			Db	386	Db 120 PNDELHRYMWPCTDYS-KGISS--FPLMLVQLTFRPGWSIGAQHHVCDRMSHFEEN 176	
			Qy	386	Qy 170 KSMGAVFRC-----CRENITH----PSLSEAAALLFPP--RDLPLPEXYDQNEALWFA 215	
			Db	386	Db 177 NSVARIANGLPLAEPDYLHLCPRNPPQIYTHSQEPFTVSLPKFLID-----228	
			Qy	386	Qy 216 GKKVATR-FVPGWKAISIODEBAK-SESVPKPSRTHATVGFMLKHLIAASRAITSGTTS 273	
			Db	386	Db 229 GRTKSQTLKFLKSERQINTLKKQDLSWNNTTRLSLYEVAGHWRV-VSKARGI-SDHDE 286	
			Qy	386	Qy 274 TRLSIAGAVKWRTRRANETYLIDNATGFLFWQAIILSHTEPFISDRLKCDUWNLNG 333	
			Db	386	Db 287 IKL---IMPVDGSRSRINNPPSLPKGCGC-----VVFLLAVCTATV3DL-ACNFTDTAG 335	
			Qy	386	Qy 334 SYRQCNQGYFETPKGKGYGRMCYELDFQRTMSSMEAPDI-----YLFS 378	
			Db	386	Db 336 RYD-----EAIKGLD----DYL- RSDAIDHTESRPDLPYPMGSPKETLIPRTL 381	
			Qy	386	Qy 379 SWTNF-ENPLDFGWGRTSWIGTGGKIBASCKFILLYPTQCGSG-IEAWVNLEEKRMAM 436	
			Db	386	Db 382 SWCRIPYQAMDFGWGNPTEFGTSNIFYDGQC---FLPSQNGDSMTAINLFSHLSHP 438	
			Qy	386	RESULT 9	
			Db	386	anthranilate N-benzoyltransferase (EC 2.3.1.144) (clone pchbct1a) - clove pink (fragment)	
			Qy	386	MEPADIVLFLSSWTFNPLDFG-WGRTSW--IIVAGKLLASCRELIVPQCGS-GIE 423	
			Db	386	HTTP:PSDLKUDLQDLYNLLNGSYKQNCDFET-----FKGKEGYGRMCYELDFQRTMSS 367	
			Qy	386	314 -TANELTKPKEFLARLKEAKSSVTEYMRSLADLMVTKRPSFSDDGAYL-----369	
			Db	386	319 -TANELTKPKEFLARLKEAKSSVTEYMRSLADLMVTKRPSFSDDGAYL-----369	
			Qy	386	320 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Db	386	321 HTTP:PSDLKUDLQDLYNLLNGSYKQNCDFET-----FKGKEGYGRMCYELDFQRTMSS 367	
			Qy	386	322 -TANELTKPKEFLARLKEAKSSVTEYMRSLADLMVTKRPSFSDDGAYL-----369	
			Db	386	323 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Qy	386	324 HTTP:PSDLKUDLQDLYNLLNGSYKQNCDFET-----FKGKEGYGRMCYELDFQRTMSS 367	
			Db	386	325 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Qy	386	326 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Db	386	327 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Qy	386	328 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Db	386	329 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Qy	386	330 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Db	386	331 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Qy	386	332 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Db	386	333 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Qy	386	334 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Db	386	335 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Qy	386	336 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Db	386	337 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Qy	386	338 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	
			Db	386	339 SPLWYRTHAIR--EDQDKERLIL--IIVARSCKKNNPDPGRTGYNAPPVIA-----318	

R; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Ohayam, L.; Fallon, L.; euss, D.; Nieman, W.C.; McHale, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, N. *Nature* 402, 761-768, 1999

R; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Ohayam, L.; Fallon, L.; euss, D.; Nieman, W.C.; McHale, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, N. *Nature* 402, 761-768, 1999

A:Residues: 1-464 <STO>
 A:Cross references: GB:AAE005172; PID:96910569; PID:AAFF31274.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F9L11.9

A:Map position: 1

Query Match 9.6%; Score 228; DB 2; Length 464;
 Best Local Similarity 22.8%; Pred. No. 4.5e-11;
 Matches 108; Conservative 94; Mismatches 171; Indels 100; Gaps 26;

Db 18 TSSTPLOPXTLTL-LDQIITPAVPIVFFYPTIDHDNLPOTLADRQASSETLTLYP 75

Db 22 TAALPLQDAWLPMSMDLILPPVQ1SVCYKXKPRHLSVAT--LKSASAEALVSYA 78

Qy 76 LSGR-VKANLYIDDF--ERGGVPELEARVNCDMDFDLRLRKIECLNEFPIKPSMEEA 131

Db 79 FAGELVKNSGSEGEPEITLCCNNGVDPFLPEAVADVELRELNIDPDESIAKLVPKK 130

Qy 132 DERYPLLGQVNVFDSG-JAIGVSYSHKLIDGFTADCFLKSGAVF--GCBE 181

Db 131 -XKSVIAQTYTQLRKCGSTIVGCFDFDHTADAFNMNMFYWSWATISRPVAPISSVSPFR 188

Qy 182 NIIHPSL---SEANALPPRDD--LPEKVDQMEALWPAKREVTRREVFGVKAISI 234

Db 189 S1LNPRRPLIIDSSTDKMMYMPVTSIPLPQETNDLNI---LTSRIVYYIKENALDL 242

Qy 235 QDEAKSESVKPK---SRVHAUTFLWKGHLIA--ASRAITSGTISTRULSAQAVNL 287

Db 243 QTLA-SGSSPKTGYCQRTKLESFSFLWKL-LVAKHTGRDLSVSKNS-KNGI---VVDGFR 296

Qy 288 RMMNETVLLONATGNNLFWNAQATLLESTHTTPBISDLKLCDLVNLNLNSVHQNGDYFETPK 347

Db 297 RL-MEKEDNTYFGN----VLSLPGGGGSDLI-IDKPLSHNTNEVER----FLEEA 343

Qy 348 GKEGYGRMCBEYLDFORTMSSME-----PADDIYLFSTMTNF-FNPLDFGNG--- 392

Db 344 TKEHFLNLIDWETHRPIAVSRYSTGTNDGPA---FVWSGSEFSPVNKVDFGWGLPVF 400

Qy 393 --RTSWSWIGAKTESASCKFILVPTOCGSGTIAWV--NLEBEKMMIQLD 439

Db 401 GSTHFPWEGSSG-----YVMPMPSVVDGNGDWVVTIHLTKGQLKPIEE 445

RESULT 13

T01056

hypothetical protein YUP8H12R.39 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-eared cress)

C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 22-Oct-1999

C:Accession: T01056 R: Theologis, A.; Vysotskaya, V.S.; Osborne, B.I.; Schwartz, J.R.; Federspiel, N.A.; Kwan

Offerer, P.; Davis, R.W. submitted to the EMBL Data Library, May 1998

A:Description: Arabidopsis thaliana chromosome 1 YAC YUP8H12R sequence.

A:Accession: T01056

A:Macolecule type: DNA

A:Status: translated from GB/EMBL/DBJ

A:Residues: 1-572 <THE>

A:Cross-references: EMBL:AC002986; PID:92494106; PID:93152598; GSPDB:GN00059; ATSP:XUP8H

C:Genetics:
 A:Gene: ATSP:XUP8H12R.39

A:Map position: 1

A:Introns: 71/2; 98/1; 257/3

Query Match 9.4%; Score 222; DB 2; Length 572;

Best Local Similarity 22.9%; Pred. No. 1.9e-10;

Matches 115; Conservative 86; Mismatches 159; Indels 142; Gaps 28;

Query Match 9.4%; Score 222; DB 2; Length 572;

Best Local Similarity 22.9%; Pred. No. 1.9e-10;

Matches 115; Conservative 86; Mismatches 159; Indels 142; Gaps 28;

Qy 6 VSINSKHTKPKSTSPLQPKYLTL--LDDQTPPAYPIVFFYPTIDHDENLPCTLA--D 61

Db 130 VTTTRKQVI--TANLPLCTHWEFLSNLILPILNIVYCFKPKLHTN--TVAYET 183

Qy 62 LROALSETLTLYYPLPSGRYKNN-----LYDDDFEGGPYLEARVNCDMDFDLRKIEC 115

Db 184 LKTLAETLVYYAFAGEIWTNTPTGPETLICNN -RGDFVFEAGADVEIRELYDPPES 241

Qy 116 LNEFPKPKPSMEEAISDEMPYLLQVNWFDG-IAIGVSVSHKLIDGFTADCFLKSGA 174

Db 242 IAKLYPIK-----RKHGVIATQVQIKCGSIVVGCFTFDHRVADAYEMMNPFLSVAE 291

Qy 175 VFR----GC-----RENTIHPSEARL---LPPRDDL-PEKYQDMEAL 212

Db 292 ISRSWVPIPSVPSFRSLNPRVMDSESIDQYIMPTSLPPQETTINFL----- 344

Qy 213 WFGKVKVATRREVFGVKAISIODEAKSEVKPSRVHAVTGF-WKHLIAASRALSGT 272

Db 345 -----LASRIYYIKKANALQELQTASSSSNGKRTKLEFSFSAF-LVAEHAAKDPVPI 397

Qy 52 ---DPNLPQTFLDRLRQALSETLTLYYPLPSGRYKNN-----LYDDDFEGGPYLEARVNCDMFT 105

Db 60 MYFDANI-----LIEALSKALVPPYPMGRKLKINGDREIDCNDEGALVEAESSHVL 113

Qy 106 DFLRLRKIECLNE-FVPKPKPSMEEAISDEMPYLLQVNWFD-SGIAIGVSVSHKLIDGFT 163

Db 114 DFGDPFRNDLHLRYMVPCTDYS-KGSS -FPLMVMQLIRFRCGVSVIGAQHHHVCDGM 170

Qy 164 TADCFLKPKSTSPLQPKYLTL--CRENITH-----PSLSEARLIPPP-RDLPKPYQDM 209

Db 171 AHFBNNSWARIKCLPAPLEPVHDYHLRPRNPSPQKXSHQSPEPFBPSLPNLLD-- 228

Qy 210 BALWFGKVKATR-FVFGVKAISIODEAKSEVKPSRVHAVTGF-WKHLIAASRAL 267

Db 229 -----GKTKNSCCLFLSREQINTLKOXKLDLSNNTTRLSTYEVAAHWSR-VSKARGL 281

Qy 268 TSGTSTRLSTIAQAVNLTRPMNETVLDNATGTFWVAQALESHIPTPEISDLK--L 324

Db 282 -SDHBEIKL---IMPVDRGSRINNESSLPGYCGN -----VVFLAVCATVGDLSCNPL 330

Db 325 CDLYNLNGSSEVKOONGDYPEFTFKGRGGGRMCYCEVLFDFPMSM-----EPADPY-- 375

Db 331 TDASKWNEALKGLDDYIWS-----AIDHTESKPGPLVPMGMSPEKPTLYPNV 378

Qy 376 LFSSQTWF--NPFLDGFSGRTSWIGVAGKLESASCKFPLVPTOCGSG 421

Db 379 LVNSKSGRIPYQAMDFGWGMSPTFFGTSNIPYDQGC---FLPSRDEDG 422

Db 345 -----ASRAITSGTISTRULSAQAVNL 287

Db 359 QTLA-SGSSPKTGYCQRTKLESFSFLWKL-LVAKHTGRDLSVSKNS-KNGI---VVDGFR 296

Db 373 -----SRVHAUTFLWKGHLIA--ASRAITSGTISTRULSAQAVNL 287

Db 381 QDEAKSESVKPK---SRVHAUTFLWKGHLIA--ASRAITSGTISTRULSAQAVNL 287

Db 395 -----RTSWSWIGAKTESASCKFILVPTOCGSGTIAWV--NLEBEKMMIQLD 439

Db 403 GSTHFPWEGSSG-----YVMPMPSVVDGNGDWVVTIHLTKGQLKPIEE 445

Db 411 -----ASRAITSGTISTRULSAQAVNL 287

Db 429 -----SRVHAUTFLWKGHLIA--ASRAITSGTISTRULSAQAVNL 287

Db 461 -----ASRAITSGTISTRULSAQAVNL 287

Db 475 -----ASRAITSGTISTRULSAQAVNL 287

Db 489 -----ASRAITSGTISTRULSAQAVNL 287

Db 503 -----ASRAITSGTISTRULSAQAVNL 287

Db 527 -----ASRAITSGTISTRULSAQAVNL 287

Db 541 -----ASRAITSGTISTRULSAQAVNL 287

Db 565 -----ASRAITSGTISTRULSAQAVNL 287

Db 579 -----ASRAITSGTISTRULSAQAVNL 287

Db 593 -----ASRAITSGTISTRULSAQAVNL 287

Db 607 -----ASRAITSGTISTRULSAQAVNL 287

Db 611 -----ASRAITSGTISTRULSAQAVNL 287

Db 625 -----ASRAITSGTISTRULSAQAVNL 287

Db 635 -----ASRAITSGTISTRULSAQAVNL 287

Db 649 -----ASRAITSGTISTRULSAQAVNL 287

Db 663 -----ASRAITSGTISTRULSAQAVNL 287

Db 677 -----ASRAITSGTISTRULSAQAVNL 287

Db 699 -----ASRAITSGTISTRULSAQAVNL 287

Db 713 -----ASRAITSGTISTRULSAQAVNL 287

Db 727 -----ASRAITSGTISTRULSAQAVNL 287

Db 759 -----ASRAITSGTISTRULSAQAVNL 287

Db 773 -----ASRAITSGTISTRULSAQAVNL 287

Db 795 -----ASRAITSGTISTRULSAQAVNL 287

Db 817 -----ASRAITSGTISTRULSAQAVNL 287

Db 832 -----ASRAITSGTISTRULSAQAVNL 287

Db 860 -----ASRAITSGTISTRULSAQAVNL 287

Db 876 -----ASRAITSGTISTRULSAQAVNL 287

Db 894 -----ASRAITSGTISTRULSAQAVNL 287

Db 912 -----ASRAITSGTISTRULSAQAVNL 287

Db 936 -----ASRAITSGTISTRULSAQAVNL 287

Db 948 -----ASRAITSGTISTRULSAQAVNL 287

Db 962 -----ASRAITSGTISTRULSAQAVNL 287

Db 976 -----ASRAITSGTISTRULSAQAVNL 287

Db 988 -----ASRAITSGTISTRULSAQAVNL 287

Db 1002 -----ASRAITSGTISTRULSAQAVNL 287

Db 1016 -----ASRAITSGTISTRULSAQAVNL 287

Db 1040 -----ASRAITSGTISTRULSAQAVNL 287

Db 1054 -----ASRAITSGTISTRULSAQAVNL 287

Db 1068 -----ASRAITSGTISTRULSAQAVNL 287

Db 1082 -----ASRAITSGTISTRULSAQAVNL 287

Db 1096 -----ASRAITSGTISTRULSAQAVNL 287

Db 1110 -----ASRAITSGTISTRULSAQAVNL 287

Db 1134 -----ASRAITSGTISTRULSAQAVNL 287

Db 1162 -----ASRAITSGTISTRULSAQAVNL 287

Db 1200 -----ASRAITSGTISTRULSAQAVNL 287

Db 1224 -----ASRAITSGTISTRULSAQAVNL 287

Db 1260 -----ASRAITSGTISTRULSAQAVNL 287

Db 1300 -----ASRAITSGTISTRULSAQAVNL 287

Db 1360 -----ASRAITSGTISTRULSAQAVNL 287

Db 1380 -----ASRAITSGTISTRULSAQAVNL 287

Db 1400 -----ASRAITSGTISTRULSAQAVNL 287

Db 1436 -----ASRAITSGTISTRULSAQAVNL 287

Db 1462 -----ASRAITSGTISTRULSAQAVNL 287

Db 1500 -----ASRAITSGTISTRULSAQAVNL 287

Db 1540 -----ASRAITSGTISTRULSAQAVNL 287

Db 1580 -----ASRAITSGTISTRULSAQAVNL 287

Db 1640 -----ASRAITSGTISTRULSAQAVNL 287

Db 1680 -----ASRAITSGTISTRULSAQAVNL 287

Db 1720 -----ASRAITSGTISTRULSAQAVNL 287

Db 1780 -----ASRAITSGTISTRULSAQAVNL 287

Db 1840 -----ASRAITSGTISTRULSAQAVNL 287

Db 1880 -----ASRAITSGTISTRULSAQAVNL 287

Db 1920 -----ASRAITSGTISTRULSAQAVNL 287

Db 1980 -----ASRAITSGTISTRULSAQAVNL 287

Db 2020 -----ASRAITSGTISTRULSAQAVNL 287

Db 2080 -----ASRAITSGTISTRULSAQAVNL 287

Db 2120 -----ASRAITSGTISTRULSAQAVNL 287

Db 2160 -----ASRAITSGTISTRULSAQAVNL 287

Db 2200 -----ASRAITSGTISTRULSAQAVNL 287

Db 2280 -----ASRAITSGTISTRULSAQAVNL 287

Db 2320 -----ASRAITSGTISTRULSAQAVNL 287

Db 2380 -----ASRAITSGTISTRULSAQAVNL 287

Db 2420 -----ASRAITSGTISTRULSAQAVNL 287

Db 2500 -----ASRAITSGTISTRULSAQAVNL 287

Db 2580 -----ASRAITSGTISTRULSAQAVNL 287

Db 2640 -----ASRAITSGTISTRULSAQAVNL 287

Db 2700 -----ASRAITSGTISTRULSAQAVNL 287

Db 2760 -----ASRAITSGTISTRULSAQAVNL 287

Db 2820 -----ASRAITSGTISTRULSAQAVNL 287

Db 2900 -----ASRAITSGTISTRULSAQAVNL 287

Db 2980 -----ASRAITSGTISTRULSAQAVNL 287

Db 3020 -----ASRAITSGTISTRULSAQAVNL 287

Db 3100 -----ASRAITSGTISTRULSAQAVNL 287

Db 3140 -----ASRAITSGTISTRULSAQAVNL 287

Db 3200 -----ASRAITSGTISTRULSAQAVNL 287

Db 3240 -----ASRAITSGTISTRULSAQAVNL 287

Db 3300 -----ASRAITSGTISTRULSAQAVNL 287

Db 3340 -----ASRAITSGTISTRULSAQAVNL 287

Db 3400 -----ASRAITSGTISTRULSAQAVNL 287

Db 3440 -----ASRAITSGTISTRULSAQAVNL 287

Db 3500 -----ASRAITSGTISTRULSAQAVNL 287

Db 3540 -----ASRAITSGTISTRULSAQAVNL 287

Db 3600 -----ASRAITSGTISTRULSAQAVNL 287

Db 3640 -----ASRAITSGTISTRULSAQAVNL 287

Db 3700 -----ASRAITSGTISTRULSAQAVNL 287

Db 3740 -----ASRAITSGTISTRULSAQAVNL 287

Db 3800 -----ASRAITSGTISTRULSAQAVNL 287

Db 3840 -----ASRAITSGTISTRULSAQAVNL 287

Db 3900 -----ASRAITSGTISTRULSAQAVNL 287

Db 3950 -----ASRAITSGTISTRULSAQAVNL 287

Db 4010 -----ASRAITSGTISTRULSAQAVNL 287

Db 4070 -----ASRAITSGTISTRULSAQAVNL 287

Db 4110 -----ASRAITSGTISTRULSAQAVNL 287

Db 4150 -----ASRAITSGTISTRULSAQAVNL 287

Db 4210 -----ASRAITSGTISTRULSAQAVNL 287

Db 4250 -----ASRAITSGTISTRULSAQAVNL 287

Db 4310 -----ASRAITSGTISTRULSAQAVNL 287

Db 4350 -----ASRAITSGTISTRULSAQAVNL 287

Db 4410 -----ASRAITSGTISTRULSAQAVNL 287

Db 4450 -----ASRAITSGTISTRULSAQAVNL 287

Db 4500 -----ASRAITSGTISTRULSAQAVNL 287

Db 4560 -----ASRAITSGTISTRULSAQAVNL 287

Db 4600 -----ASRAITSGTISTRULSAQAVNL 287

Search completed: July 6, 2004, 13:39:30
Job time : 21.1722 secs

QY	179 ------CRENTIHPSTLSEAA11FFPPRDLPLPEKYYDQMLWAGKKVATRFF 224	QY	106 DELURKITECLNFNFVPIKP-----FSGMEAISD-ERYPLIIGYQVNVD-F-SGIAIGVSVSH 157
Db	186 VPKPSLEPIWNRELLINPDPHLQLNQFDSCCPP-----PMLBELQAS-----F 229	Db	109 D----NELSVTLGDPDSDNSPQQLLPSLPLDNLVQVTRFTCGFVVSFSFH 164
QY	225 VFGYKAISSIQDEAKSESVSPKPSRPHAVTGFELKHLIASRAL-TSGTISTRLSIAKQV 283	QY	158 KLDGCTADCEFLXSGAVFRGCREMILHPSLSEALLFPPRDLPLPEKYYDQMLWFF-- 214
Db	230 VINYDTEYKMCYMEBCNEFCSSFEVVAALVV--IAKTKALQPHPEVNLKF--AM 283	Db	165 GVDCTGRRGAQFLKGIAEMRG-----EYKLSELPIMNRELVKLUDDPKVLFQFHF 213
QY	284 NLRTRMNMETVLDNATGNEFLWAQILELSTHTEPSDLKLUKCDLYNLNGS-----VK 336	QY	215 -----AGKQVATPRRPSVQKAISSIQDEAKSSEVPKSRHAVTGFELKHLIASRAL 267
Db	284 DLRLKFN_PPLPNQYYGNAGTATA-----MDNFDLNLSSLLPAlMIK 327	Db	214 EFLRAPSIVEKIVQTYFIDFETINYKQSQVMECKEFCSSFEVASYAMTA--IARTRAF 270
QY	337 QCNGDVFETFKGKREGYGRMCYELDFQRTMSSMEPAFDYIYPTQ 417	QY	268 TSGTISTRLSIAQAVNLRTTRMMETVLD----NATGNLFWAACILELSHTTPEIISDL 322
Db	328 KAKADL----KDNYSR-SRVTINPYSLDVNKKSNDNIALSDMRLGFTYEADFGWGGPL 380	Db	271 QIP-----ESEVTKLLEGMDRNSFRPPLPGYYGNSIGTACV----- 309
QY	396 WIGVAGKTESA---SCKPTILVYPTQ 417	QY	323 KLCDLVNLINGS-----WQQCNGDVFETFKGKREGYGRMCYELDFQRTMSSMREPA----- 371
Db	381 NVSSQRLNGLPMSTFLYLPK 405	Db	310 -DNYQDLSLSSLLRMAITKSKVSLNDNFKSR-----AVVPESELDV 351
RESULT 2			
TSAT_TAXCH	STANDARD;	PRT;	439 AA.
AC	QBS9G6;		
DT	28-FEB-2003 (Rel. 41, Created)	DT	372 -DIXYLFSSTWTF-FNPLDFGNGRTSMIG-VAGKIESASCK-FILVYPTQ-CGSIGIE 423
DT	28-FEB-2003 (Rel. 41, Last sequence update)	DT	352 NMNENVVAFADWSSRLGDEVDFGNGNAVSVPNCQQQCELAMONYFLFKPSKNKPDGK 411
DE	Taxadien-5-alpha-ol-O-acetyltransferase (EC 2.3.1.162) (Taxa-	QY	424 AWNLBEKRM 433
DE	4(20)-11(12)-dien-5-alpha-ol-O-acetyltransferase) (Taxadienol-	Db	412 ILMFLPFLSHM 421
DB	acetyltransferase).		
OS	Taxus chinensis (Chinese Yew).	ID	DBAT TAXCH STANDARD: PRT; 440 AA.
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;	AC	Q9MGF2; (Rel. 41, Created)
OC	Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.	DT	28-FEB-2003 (Rel. 41, Last annotation update)
RN	NCBI_TaxID=29808;	DT	28-FEB-2003 (Rel. 41, Last annotation update)
RP	SEQUENCE FROM N.A.	DB	10-deacetylbbaccatin III 10-O-acetyltransferase (EC 2.3.1.167) (DBAT).
RA	Wang W., Song S., Cheng K.	OS	Taxus cuspidata (Japanese Yew).
RT	"Cloning of a tax-4(20),11(12)-dien-5-alpha-ol-O-acetyltransferase	OC	Bukaryota; Viridiplantae; Streptophyta; Tracheophyta;
RT	cDNA from Taxus chinensis and functional expression in Escherichia	OC	Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
RT	coli."	ON	[1]
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.	RN	SEQUENCE FROM N.A., AND CHARACTERIZATION.
CC	- CATALYTIC ACTIVITY: Acetyl-CoA + taxa-4(20),11-dien-5-alpha-ol =	RP	Proc. Natl. Acad. Sci. U.S.A. 97:583-587(2000).
CC	COA + taxa-4(20),11-dien-5-alpha-ol acete.	RX	LINE 2:010515; PubMed=10539122;
CC	- PATHWAY: Taxol biosynthesis; third step.	RA	Walker K., Croteau R.B.;
CC	- SIMILARITY: Belongs to the plant acetyltransferase family.	RT	"Molecular cloning of a 10-deacetylbbaccatin III-10-O-acetyl
CC		RT	transferase CDNA from Taxus and functional expression in Escherichia
CC		RT	coli."
CC		RL	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC		CC	the European Bioinformatics Institute. There are no restrictions on its
CC		CC	use by non-profit institutions as long as its content is in no
CC		CC	way modified and this statement is not removed. Usage by and for commercial
CC		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).
CC		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC		CC	the European Bioinformatics Institute. There are no restrictions on its
CC		CC	use by non-profit institutions as long as its content is in no
CC		CC	way modified and this statement is not removed. Usage by and for commercial
CC		CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).
DR	EMBL; AY078205; AY078254.1; -.	CC	CC
DR	InterPro; IPR003480; Transferase.	DR	DR
DR	PF02458; Transferase; 1.	DR	DR
KW	TAXOL BIOSYNTHESIS; Transferase; Acyltransferase.	PF02458; Transferase; 1.	DR
SEQUENCE	49183 MW;	SEQUENCE 440 AA;	KW
SEQUENCE	59EC6EE807989C843 CRC64;	SEQUENCE 440 AA;	SEQUENCE 440 AA;
QY	56 PQTL-AD---LROALSETLTLVYPLPSLRKGVF---NNLYDDFEEGVYPLLEARVNCDMT 105	6.3%; Score 148.5; DB 1; Length 440;	Query Match 6.3%; Score 148.5; DB 1; Length 440;
Db	54 SPTMVSADPAKLIREALAKLIVVYPPFAGRLRETEGDLVECTGEAMLEA---MA 108	Best Local Similarity 23.6%; Pred. No. 6.8e-05;	Best Local Similarity 23.6%; Pred. No. 6.8e-05;

Matches 99; Conservative 62; Mismatches 158; Indels 101; Gaps 20;

Query 14 KPKSTSTSPLOQYKLTLDQDTPPAYPIVF---FPPITDIDENLQTLDPAYLSET 69
 DB 17 VAPSQPS-PAFQLQLSTDNL---PGVBNINTLLVNASDVSDBAKV---IRQALSKV 71

Query 73 LTYYPLSGRK---NNLYIDDEEEGVPYLEARVNCDMTFLURKIECLNEFVPIKPF 125
 DB 72 LYVSPFAGRLLRKENGDLVERTGEAGALFVAMADTD----LSTVGDLDYSP---- 121

Query 126 SMEAI----SDERYPLIGQVQVNF---SGIAIGYVSHSKLIDGTAFCFLKSNGAV 175
 DB 122 SLEQLLFCLPDPDTDIDHLPV-VQVTRTPCGFVWVYSPCFCGQFLGAMGEM 180

Query 176 FRCRENTIHPSEAAI---IIPPRDOLP---EYVDOMEAALWFGAKRWRFFVFGYKA 230
 DB 181 ARG---EIKPS-SEP1WKEELIKPDELYRPFQYHQLICPPSTFCGKTIWQSSLVITSET 235

Query 231 ISSIQEAKSESVPKPSRVSRYHATVGEMLKHLIASRAL-TSGTTISTRULSIARQAVNLRTRM 289
 DB 236 INCIKOCILRESEKFCSAFEVSLAN---IARTALQJQPHSENVRLLF---AMDWERKLF 289

Query 290 N-----METVLDNATGNLFWWQAQALESHTTPEISDLKLCDLYNLN 332
 DB 290 NPLSKGYYGNGFVGTVCAMDNDVDRDLSSEL---RVVRIKGRKAVSLSNEHTSTIVTRS 346

Query 333 GSYVQCNEDYFETFGKREGYGRMCYLDFTQTMSSMEPAPDLYLPSSTTFNPLDQG 392
 DB 347 GSDESINYENITVSGFDRRLG----FDEDEVDFSGW 376

RESULT 4
 ID TSAT TAXCU STANDARD; PRT: 439 AA.
 AC Q9M6F0;
 DT 28-FEB-2003 (Rel. 4.1, Created)
 DT 28-FEB-2003 (Rel. 4.1, Last sequence update)
 DT 28-FEB-2003 (Rel. 4.1, Last annotation update)
 DE 4 (20).11 (12)-dien-5alpha-ol-0-acetyltransferase (EC 2.3.1.162) (Taxa-
 DE acetyltransferase).
 DE TAT.
 OS Taxus cuspidata (Japanese yew).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
 CC NCBI_TaxID=99806;

CC RN 11] SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.

CC RX MEDLINE=2012B77; Published=1066320;
 CC RA Walker K., Schoendorf A., Croteau R.B.;
 CC RT "Molecular cloning of a tax-4(20),11(12)-dien-5alpha-ol-0-acetyl-
 CC transferase cDNA from Taxus and functional expression in Escherichia
 CC coli."
 CC RL Arch. Biochem. Biophys. 374:371-380 (2000).
 CC -I CATALYTIC ACTIVITY: Acetyl-CoA + tax-4(20),11-dien-5-alpha-ol =
 CC COA + tax-4(20)-11-dien-5-alpha-yl acetate.
 CC -I SIMILARITY: Taxol biosynthesis; third step.
 CC -I SIMILARITY: Belongs to the plant acetyltransferase family.

CC This SNSS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license@ebi.sib.ac.uk/announce/
 CC or send an email to license@ebi.sib.ac.uk.
 CC DR EMBL; AF09130; AA342454.1; -.
 CC DR PIB: T52321; T52321;
 CC DR InterPro; IPR03480; Transferase.
 CC DR Pfam; PF00458; Transferase_1.
 CC XW Taxol biosynthesis; Transferase; Acyltransferase.
 CC SEQUENCE 439 AA; 49079 MW;

Query Match 6.1%; Score 144; DB 1; Length 439;
 Best Local Similarity 21.0%; Pred. No. 0.00016;
 Matches 78; Mismatches 188; Indels 118; Gaps 22;

Query 1 MEKIEVSIN-SKHTKIPSTSSTPLQPKYLTLQDQITPPAYPPIVFYPITDHDENLQQT 58
 DB 1 MERTDLYHNLIEKVMGDS-EPPLPKTQLQSSIDNL---PGVGRSISFARLLITYNASPSPMT 57

Query 59 LA----DLRQALSETLTLYPLSPGRK---NNLYIDDEEEGVPYLEARVNCDMTFL 109
 DB 58 ISADPAKPIEALAKILVYVPPAGRLRETENGDLVEFVGTGEMFILEA----MAD--- 109

Query 110 LRKIECLNEFVPIKPK----FSMEAISD-ERYPLIGQVQVNTD-SGIAIGYVSHKLID 161

DB 110 -NEASVLSLDEDNSNPSPQOLLSPLDINFQDLSLWVQVTRRTGCGWVGVSPFHGVCDF 168

Query 112 GGTADCFELKSWGAVERGCRENTIHPSEAAI.FPPRDLPLPEKVDOMEALWF---- 214

DB 116 169 GRGAQFLGLAEMARG-----EVKQSLEPITWNRELVFLDKYLQFEEFEFLR 217

Query 121 -AGKTVATRTRVFGVKAISSTQDEAKSESVPKPSRTHAVTCFLWKLHIAARALSGT 271

DB 128 APSIVKEVQYTLIDPEFTINVKQSYMEECFKCSPEVASAMTW---IARTRAFQIP- 273

Query 162 TSTRSLTAAQVNLRTKRMNMYLTD----NATGNLFLFWWQAQELSHTPETSDLKLCD 326

DB 174 -----ESEYVXKILFGMDRNSNFPPLPSGYGNSIGTACAV----DN 311

Query 327 LYNNLGS-----VKQCKGDFETFGKEGYGRMCYLDFTQTMSSMEPA----- 371

DB 312 VQDLISGSLLRMAIMIKRSKVSLSNDNFNSR-----AVVKPSELDVNMMH 355

RESULT 5
 ID ANTA GENTR STANDARD; PRT: 469 AA.
 AC Q9ZPR8; 15-MAR-2004 (Rel. 4.3, Last sequence update)
 DT 15-MAR-2004 (Rel. 4.3, Last annotation update)
 DB 356 ENTAFA0D9RSLEDEVDFGWNAVSVSPVQDQSLAMQNYFLPKPSRNKPDK1KLMF 415

Query 372 PDYLFSSKTNF_FNPDLFGWGRITSWIGTAGKIESASCK---FILVPTQ-CGSSIEATWN 427

DB 428 LEEFKM 433

Query 416 LPISKM 421

DB 416 LPISKM 421

RN SEQUENCE FROM N.A.; SEQUENCE OF 68-89; 249-262; 395-404 AND 442-459,
 RP CHARACTERIZATION, TISSUE SPECIFICITY, DEVELOPMENTAL STAGE, AND
 RP SUBCELLULAR LOCATION.

CC TISSUE=Petal; RX MEDLINE=99097837; PubMed=9881162;
 CC RC Fujiwara H., Tanaka Y., Yonekura-Sakakibara K., Fukuchi-Mizutani M.,
 CC RA Nakao M., Fukui Y., Yamaguchi M., Ashikari T., Kusumi T;
 CC RT "CONA cloning, gene expression and subcellular localization of
 CC anthocyanin 5-aromatic acyltransferase (EC 2.3.1.153) (SAT).
 CC OS Gentiana trifolia (Clustered gentian).
 CC OC Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 CC OC Lamiales; Gentianales; Gentianaceae; Gentiana.
 CC NCBI_TaxID=55190;

CC [1] RN SEQUENCE FROM N.A.; SEQUENCE OF 68-89; 249-262; 395-404 AND 442-459,
 CC RP CHARACTERIZATION, TISSUE SPECIFICITY, DEVELOPMENTAL STAGE, AND
 CC SUBCELLULAR LOCATION.

CC TISSUE=Petal; RX MEDLINE=99097837; PubMed=9881162;
 CC RC Fujiwara H., Tanaka Y., Yonekura-Sakakibara K., Fukuchi-Mizutani M.,
 CC RA Nakao M., Fukui Y., Yamaguchi M., Ashikari T., Kusumi T;
 CC RT "CONA cloning, gene expression and subcellular localization of
 CC anthocyanin 5-aromatic acyltransferase from Gentiana trifolia.";
 CC RL Plant J. 16:421-431 (1998).
 CC -I FUNCTION: Transfers hydroxycinnamic moieties to the glucosyl
 CC CC groups of anthocyanin.
 CC -I CATALYTIC ACTIVITY: Hydroxycinnamoyl-CoA + anthocyanidin-3-5-
 CC CC di glucoside = CoA + anthocyanidin-3-glucoside-5-
 CC CC hydroxycinnamoylglucoside.
 CC -I ENZYME REGULATION: Activity enhanced by manganese ions and
 CC inhibited by p-chloromercuribenzoate.

CC -i- SUBCELLULAR LOCATION: Cytoplasmic. Expressed in the outer epidermal cells of petals, but not in sepals, leaves or stems.

CC -i- DEVELOPMENTAL STAGE: Induced during petal development and then decreases rapidly with maturation of the flower.

CC -i- PTM: The N-terminus is blocked.

CC -i- MISCELLANEOUS: Cannot use malonyl-CoA as a donor.

CC -i- SIMILARITY: Belongs to the plant acetyltransferase family.

CC -i- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sb-sib.ch).

CC -i- DR: EMBL; AB010708; BAA74428; 1;

DR: PR02448; Transferase; 1;

KW: Transferase; Acyltransferase.

DR: PR02448; Transferase; 1;

DR: 52739 MW;

SEQUENCE: 469 AA;

EDFFB0B5E5963ACB CRC64;

Query Match Score: 6.0%; Score: 143; DB: 1; Length: 469;

Best Local Similarity: 19.3%; Pred. No.: 0.00022;

Matches: 96; Conservative: 83; Mismatches: 186; Indels: 132; Gaps: 23;

Qy 23 LQPYKLTLDQLTPPA-----YVPIVPP-----YVITDHDNLPLQT-----L 59

Db 4 IOMYKULEKCOVTPPSDTDVEILSLPVTFEDIPWTHNMQSLFYDPYPRTHFLDTV 63

Qy 60 ADLRLQALSETLTLYVPLSGR-----VKN-----NLYIDDFEGPYLEAR 99

Db 64 PNIKASLSTLTLYVPLSGNLLMPIKSGMPKQYSRDEGDSITLIVAESQDEDYLKHH 123

Qy 100 VNCMDTDFPLRLKIECINEFVPIKPFMSMFAISDR-YPLLGQVQNF-DSGIAIGVSYSH 157

Db 124 QLYDSND-----LHQLFYMPVIRTQMDYKVPVAVQVQTVPFRGIAVALTH 174

Qy 158 KLDGGTADCFLKSWGAVRGCRE-----NIHPSLSEAL-----LFPFRDPLPKYQDM 209

Db 175 SIADAKSFYMFNTNAWAYINKFQKDADLLS-PSFDRSIIKDLGLEYETWNEMQDL 233

Qy 210 BALMFAGK-----VATREFFVGKAIISSIQDEAKSESVPKPS-RHAVT---GFLWKHL 260

Db 234 EMFMSRGSKRPREKFNQYRATVYLSAETIQLKNULNLRGSEPTIRVTFMTCSYVWTCM 293

Qy 261 IAAARSLATSTGTST-----RLSTAAQAVNLRMNMETVLDNATGULFWMQAILELSH 314

Db 294 VKSQDDVWSESSRDENELEYFSTADRGILT---PPCPNPFGNCL---ASCVKATH 347

Qy 315 TTPRISDLKCLDLYNLNLNSVK-----OCMSDYFETFKGKEGYGMC 356

Db 348 -KELVGDKSLVATAAIGRAEKRHLNEKGVLADAKTILSENGIPSRSRFLGITG-----401

Qy 357 EYLDQFRTMSSMPEPAPDYLSSWTFNFFAPLDFGKRTSWSVIGAKTISASCPIILVPT 416

Db 402 -----SPKFDSY-----GVDGSGKPAKFKDIT-SVVDYAEIYVI-QSR 437

Qy 417 QCGGSIIEAWVNLLEEFKM 433

Db 438 DFEKGVEIGVSLPKHM 454

RESULT 6

DBST TAXON ID DBBT TAXON STANDARD; PRT; 440 AA.

AC Q9PPW3; PRT; 440 AA.

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE 2-alpha-hydroxytaxane 2-O-benzoyltransferase (EC 2.3.1.166) (TBT) (2-Debenzoyl-7,13-diacetylbaccatin III-2-O-benzoyl transferase) (DBBT).

OS Taxus cuspidata (Japanese Yew).

OC Bakterien; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.

NCBI TAXID=99806;

RN [1]

RP SEQUENCE FROM N.A.

MEDLINE=2055551; PubMed=11095755;

RA Walker K, Croteau R.B.;

RT "Taxol biosynthesis: molecular cloning of a benzoyl-CoA:taxane 2alpha-O-benzoyltransferase cDNA from Taxus and functional expression in Escherichia coli";

RL Proc. Natl. Acad. Sci. U.S.A. 97:13591-13596 (2000).

-i- FUNCTION: Catalyzes the conversion of 2-debenzoyl-7,13-diacetylbaccatin III, a semisynthetic substrate, to 7,13-diacetylbaccatin III.

CC -i- CATALYTIC ACTIVITY: Benzoyl-CoA + 10-deacetyl-baccatin III. IIII = CoA + 10-deacetyl-baccatin III.

CC -i- PRMTWAY: Taxol biosynthesis.

CC -i- SIMILARITY: Belongs to the plant acetyltransferase family.

CC -i- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sb-sib.ch).

CC -i- DR: AP297618; AAC38049.1; -

DR: InterPro; IPMC0340; Transferase.

DR: PR02458; Transferase; 1.

KW: Taxol biosynthesis; Transferase; Acyltransferase.

SEQUENCE: 440 AA; 50089 MW; AAC92B403A605A CRC64;

Query Match Score: 6.0%; Score: 142.5; DB: 1; Length: 440;

Best Local Similarity: 21.4%; Pred. No.: 0.00022;

Mismatches: 52; Indels: 135; Gaps: 18;

Matches: 93; Conservative: 52; Mismatches: 155; Indels: 135; Gaps: 18;

Qy 62 LQPLSETLTLYVPLSGR-----NLYIDDFEGPYLEARVNCMDTFLRERKICLN 117

Db 63 IREALSKVLYVYPPEAGRLRNTGDLVECTGEGAVFVEMADND-----LSVLDQDEN 113

Qy 118 EFPV-----KPSM-EAISDERYPLGVQVNFVFD-SGIAIGVSVSHKL-TGTTGTADECFLKSW 172

Db 114 EYDPSFOQVLENREDVNTEDLHLTVQTRFTCGFVFGTRFHSVSORGICGQLLKGK 173

Qy 173 GAVERG-----CREMILHPSLSEALLPPRDLPEKTYDQMEALWFGK 217

Db 174 GEMARGEKPSLEPIWNREMVKPDMIVLQFDHDFDFFAPLN-----LE 217

Qy 218 KVATRFRFVGKAIISSIQDEAKSSVPRSRVHATVGFQWHLTAASRALTSGTISTRIS 277

Db 218 KSIQASMVISFERANVIRCMMECKEKFSSAFETVVALW-----LARTKSFRRP 267

Qy 278 IAAQAVNLTRMNNETVLD-----NATGNEFWAQAILEL-SHTKPEISDLKLKDLY 328

Db 268 -PNVYKXKIFFPIDMRNSFDSPLPGYNAIGN-----ACAMD 304

Qy 329 N-----LINGS-----VKQCGNDYFETPKGKEGYGMCYLDPORTMSSMPEPDIYLF 378

Db 305 NVRDQDNGSLLYALMLKIKSKFALNENPSR-----ILTKPSTLIDANMKHENVVGCG 356

Qy 379 SWTF-----FNPUDFGMGRGRTSMWVAGKLESASCKF ILVPIVOCGSIEAWVNBLERKAMMLE 437

Db 357 DWRNLGFYXADFGHGN-----VNVSPMQ-----QREHELM-----389

Qy 438 QDPHFLALASPKTLI 452

Db 390 -QNTFLFLRSAKM 403

RESULT 7

SYA_MYCPN

ID	SYA MYCPN	STANDARD;	PRT;	900 AA.	
AC	P75768;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlaRS).				
GN	ALAS OR MNM419 OR MP422.				
OS	Mycoplasma pneumoniae.				
OC	Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.				
OX	NCBI_TaxID=2104;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAT-ATCC 29342 / M129;				
RX	MEDLINE=97105885; PubMed=8948633;				
RA	Himmelreich R., Hilbert H., Plagens H., Pirkl B., Li B.-C., Herrmann R.,				
RT	*Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.;				
RL	Nucleic Acids Res. 24:4420-4449 (1996).				
CC	-!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP + diaphorinase + L-alanyl-tRNA.				
CC	-!- SUBCELLULAR LOCATION: Cyttoplasmic.				
CC	-!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ish-sib.ch/announce/ or send an email to license@ish-sib.ch).				
DR	EMBL: AECG0041; AAB96070.1; -.				
DR	PIR: S73748; S73748.				
DR	HAMAP: MF_00036; -; 1.				
DR	InterPro: IPR023218; tRNA-synt_2c.				
DR	InterPro: IPR006193; tRNA synt_Ala.				
DR	PFAM: PF01411; tRNA-synt_2c_1.				
DR	PRINTS: PR00980; TRNASYNTTHA.				
DR	TIGRFAMS: TIGR0044; alaS; 1.				
DR	PROSITE: PS00860; AA_tRNA_LIGASE_II_ALA; 1.				
KW	Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding; Complete proteome.				
KW	SEQUENCE 900 AA;				
DR	PIR: AECG0041; AAB96070.1; -.				
DR	Best Local Similarity 23.6%; Pred. No. 1.6;				
Matches	49; Conservative 32; Mismatches 80; Index 47; Gaps 7;				
Qy	3 KIRVINSKHTIKPSTSTTELOPQKLUDQLTPPAYVIFVFPITDHDFTNLQQTLLADL 62				
Db	719 KLSSETINSLHTNLNTIANPALKRQTLTQL-----NHFHLRQVITL 752				
Qy	63 R- QAISETLTLYPLSGRVKNLYLTDPEEGVPLYEARNVCDMTDFSLRKLCBLNETFV 120				
Db	763 RKCQBLNLFLKI-----TQNELKTEDPFWKQKQLAEKQBLLEKAVLASPA 815				
Qy	121 PIRFSM---EASDERYFLGQVNPW---SGIAIGSVS-----HKLIDGGTA 165				
Db	816 AVDPFLLSQYAQAVLNQHKNKLFVLLNQNSPNSPMLQGDVSXKCIQIKAHFELKGGS 875				
Qy	166 DCFLKSWGAVERGCRENIIPPSLSEAL 193				
Db	876 NNF-----FRGSFNESTVSKQAIL 896				
RESULT 8					
ID	PRLR_MELGA	STANDARD;	PRT;	831 AA.	
AC	Q91094; Q91091; Q91092;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
Query Match	4.2%	Score 99.5;	DB 1;	Length 831;	
Best Local Similarity	20.7%	Pred. No. 2.5;			
Matches	100;	Conservative 57;	Mismatches 156;	Indels 169;	Gaps 21;

QY 7 SINKHTIKPSTSSTPLQYKLTLDLQTPAYVPIFFYPTTDHFNL---PQTLADL 62
 Db 303 NITYKATNENGNSNSDPHVDTVTPQDPEANTPLEKKPKINRKPLMLTSPPLADV 362
 QY 63 RQALSETLTIVYPLSGRVKNNLYIDDFEEGVYPLEARYVNCDMTDFLRLRKIECLNEFVPI 122
 Db 363 R---SGWLTDYEL--RLKP- 402
 QY 123 KPF-----SMEAISDERYPLIGVQVNPFDSGIAIGVSYSHKLIDGGTADCFIK 170
 Db 403 KKYVQIHKCKPDKHGSWSNEWSSEVYIPIFNDFKTDIWIWLVGVLSSUJ-----CLIM 456
 QY 171 SWGAYFRGER-----ENIHPSLSEALL-----EPPRDLPEPKVY 206
 Db 457 SWTMVLRKGIRMITFLPPVPGPKXKGIDTHLETKSSEBLLSAGGCHGPFTSDCELLI 516
 QY 207 DOME----- 210
 Db 517 EYLEVEDSDHQLMWSHDSGRPSKNAKTIKRETORDSGRSODSPSLLSEKRCALPS 576
 QY 211 AIWF-----AGKK-----VATRRTFEGVKAISSTQDEAKSESVPKPSRVHAVIG 254
 Db 577 ALQIQDVRDQVAKKGRKSWSYCVASER-----KALLFNEKASASSTWAVQLPNNQPP 631
 QY 255 FLWKLHIIASAARALTSGTTSTRLSLAQAV-NLRTRMMETVLDATGNLFWQAQILELS 313
 Db 632 TEAHTSIVTEANKKTS-----TTINMNTAVLYNEERHOSLYSISSETISGGM--EKQEEEMNL 687
 QY 314 H-----TTPLISDLKLCDLYNLLNGSVKQNGDFT-----FKGKCGYGRMC 356
 Db 688 HSXKTQTTVQVRQNESNEKLFPLNHALM---DYEVHKYTRQDEBPTVLLKHKXSGKIE 743
 QY 357 EY 358
 Db 744 KY 745

RESULT 9
 BACH_BACLI STANDARD; PRT; 2607 AA.
 AC 068007;
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Bacitracin synthetase 2 (BA2) [Includes: ATP-dependent lysine synthetase; (LySA) (Lysine activase); ATP-dependent D-ornithine adenylase (D-Ora) (D-Ornithine activase); Ornithine racemase (EC 5.1.1.12)].
 DE (EC 5.1.1.12).
 GN BACB
 CS Bacillus licheniformis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 10716;
 RX MEDLINE=98089193; PubMed=9427658;
 RA Konz D., Klemens A., Schoergendorfer K., Marahiel M. A.;
 RT "The bacitracin biosynthesis operon of Bacillus licheniformis ATCC 10716: molecular characterization of three multi-modular peptide synthetases.".
 RL Chem. Biol. 4:927-937(1997).
 CC -!- FUNCTION: Activates two amino acids and incorporate a D-ornithine from its second active site into bacitracin.
 CC -!- CATALYTIC ACTIVITY: L-ornithine = D-ornithine.
 CC -!- COFACTOR: Contains 2 covalently bound phosphopantetheines (Potential).
 CC -!- PATHWAY: Cyclic peptide antibiotic bacitracin biosynthesis.
 CC -!- SUBUNIT: LARGE MULTIZYME COMPLEX OF BAI, BAA AND BA3.
 CC -!- DOMAIN: CONSISTS OF TWO MODULES WITH A C-TERMINAL EPIMERIZATION DOMAIN. EACH MODULE INCORPORATES ONE AMINO ACID INTO THE PEPTIDE PRODUCT AND CAN BE FURTHER SUBDIVIDED INTO DOMAINS RESPONSIBLE FOR SUBSTRATE ADENYLATION, THIOLATION, CONDENSATION (NOT FOR THE INITIATION MODULE), AND EPIMERIZATION (OPTIONAL), AND N

CC -!- METHYLATION (OPTIONAL). BACITRACIN IS A MIXTURE OF AT LEAST TEN CYCLIC
 CC -!- MISCELLANEOUS: DODECAPEPTIDES THAT DIFFER BY ONE OR TWO AMINO ACIDS. THE MOST
 CC -!- ABUNDANT IS BACITRACIN A, A BRANCHED CYCLIC DODECAPEPTIDE. IT
 CC -!- CONTAINS AN N-TERMINAL LINEAR PENTAPEPTIDE MOIETY (LEU-CYS-DEU-D-
 CC -!- GLU-TLE) WITH AN ISOLEUCINE-CYSTEINE THIAZOLINE CONDENSATION
 CC -!- PRODUCT AND A C-TERMINAL NEPTAPEPTIDE RING (LYS-D-ORN-TLE-D-PHE-
 CC -!- HIS-D-ASP-ASN), IN WHICH THE FREE ALPHA-CARBOXY GROUP OF THE C-
 CC -!- TERMINAL ASN IS BOUND TO THE EPSILON-AMINO GROUP OF LYSINE. IT
 CC -!- CONTAINS FOUR AMINO ACIDS IN THE D-CONFIGURATION (GLU-4, ORN-7,
 CC -!- PHE-9, AND ASP-11).
 CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
 CC -!- family.
 CC -!- SIMILARITY: Contains 2 acyl carrier domains.
 CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC -!- between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC -!- the European Bioinformatics Institute. There are no restrictions on its
 CC -!- use by non-profit institutions as long as its content is in no way
 CC -!- modified and this statement is not removed. Usage by and for commercial
 CC -!- entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC -!- or send an email to license@isb-sib.ch).
 CC -!- EMBL: AF007865; AAC06347-1; -.
 DR PIR; T31678; T31678.
 DR HSSP; P1487; JAMU.
 DR InterPro; IPR00873; AMP-bind.
 DR InterPro; IPR001242; Condensatn.
 DR InterPro; IPR006163; Pp.bind.
 DR InterPro; IPR006162; Ppantne.S.
 DR Pfam; PF00501; AMP-binding; 2.
 DR Pfam; PF00668; Condensation; 3.
 DR Pfam; PF00550; Pp-binding; 2.
 DR PRINTS; PR00154; AMPBINDING.
 DR PROSITE; PS00012; PHOSPHOPANTETHINE; 1.
 DR PROSITE; PS00455; AMP BINDING; 2.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 DR Ligase; Isomerase; Antibiotic biosynthesis; Phosphopantethine;
 KW Multi-functional enzyme; Repeat.
 KW Multi-functional enzyme; Repeat. DOMAIN 1 (LYSINE-ACTIVATING).
 FT REPEAT 535 1090 DOMAIN 1 (LYSINE-ACTIVATING).
 FT REPEAT 1547 2141 DOMAIN 2 (D-ORNITHINE-ACTIVATING).
 FT DOMAIN 1021 1088 ACYL CARRIER (ACP) 2.
 FT DOMAIN 2064 2130 ACYL CARRIER (ACP) 1.
 FT BINDING 1051 1051 PHOSPHOPANTETHINE (BY SIMILARITY).
 FT BINDING 2094 2094 PHOSPHOPANTETHINE (BY SIMILARITY).
 SQ 2607 AA; 297474 MW; FF654PAC5BBBA6F CRC64;
 Query Match score 4.2%; score 99.5%; DB 1; Length 2607;
 Best Local Similarity 17.6%; Pred. No. 13; Gaps 19;
 Matches 83; Conservative 72; Mismatches 149; Indels 167; Gaps 19;
 QY 1 MEKIEVWSINSKHTIK---PSTSSTPLQPKY-----LTLIDQLTTPAYVPI-VFF 45
 Db 985 LDQIPRTINGRADLVALEPPDRRAQAARYEARPNCTEALLSINWDLPAEQGINDFH 1044
 QY 46 YPITDHDENLPTQTLADRLQALSETIT----- 72
 Db 1045 FDIGGHSLLKAFSMAKIQSALKVETTLEKLFNHSTIQDLAYIAQKQKQYSDICKAKIK 1104
 QY 73 -YPLSLSGRVKNNLYDDEEBS-----VYPL----- 96
 Db 1105 EYPLSSAQKRLYILNQIEQGATANMPPAMKIKGRLQTDKAFAKFTLKRHESSRTSF 1164
 QY 97 -----EARVNCMDTDFLRKLI-----ECLNFVPIKPFMSMEASDERYPLIG--V 140
 Db 1165 VTINGSPVQNINEEVTFEMKYRELNCSPERMQG---IRPFEL---EXAPLRAELV 1217
 QY 141 QVNVEDSGLAIGVSYSHKLIDGTTACDFELKSWGAVERGCENIHP----- 156
 Db 1218 RVNAAHILLL-DMHHIIISGVSIGLMCEWAALY---EKEKLAPKXIQKDYSEWWQRD 1272
 QY 187 -----SLSTAALIIPPRDLPEKV-----DOMEALWPAGKCYTTRF-VFGVKAIISS 233

RESULT 11	Best Local Similarity	22.7%	Pred. No.	5.2;
MDR3_CTRIGR	Conservative	65;	Mismatches	156;
ID MDR3_CTRIGR	STANDARD;	82;	Indels	82;
AC P23174;	PRT;	1281 AA.	Gaps	22;
DT 01-NOV-1991 (Rel. 20, Created)				
DT 01-NOV-1991 (Rel. 20, Last sequence update)				
DT 16-OCT-2001 (Rel. 40, Last annotation update)				
DB Multi-drug resistance protein 3 (P-glycoprotein 3).				
GN PGY3 OR Pgp3.				
OS Cricetulus griseus (Chinese hamster).				
RT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Cricetinae; Cricetidae.				
OC NCBI TaxID=10029;				
RN [1] _				
RP SEQUENCE FROM N.A. PubMed=1685679;				
RX MEDLINE=92135836; PubMed=1685679;				
RA Endicott J.A., Sarangi F., Ling V.				
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein gene family."				
RT Gene Seq. 2:89-101(1991).				
CC -!- FUNCTION: Energy-dependent efflux pump responsible for decreased drug accumulation in multidrug-resistant cells.				
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.				
CC -!- MISCELLANEOUS: Pgp1 isoforms differ in their drug transport capabilities: Pgp1 and Pgp2 can mediate MDR, while Pgp3 apparently cannot.				
CC -!- SIMILARITY: Belongs to the ABC transporter family. MDR subfamily.				
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
DR EMBL; M60442; AAA68835.1; -.				
DR PIR; I48123; -.				
DR HSSP; P13569; I48123.				
DR InterPro; IPR003593; AAA ATPase.				
DR InterPro; IPR001140; ABC TM transport.				
DR InterPro; IPR003439; ABC transporter.				
DR Pfam; PF00664; ABC_transporter.				
DR Pfam; PF0005; ABC_tran; 2.				
DR SMART; SM0352; AAA; 2.				
DR PROSITE; PS05029; ABC_TMLF; 2.				
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.				
DR PROSITE; PS050893; ABC_TRANSPORTER_2; 2.				
KW Multi-gene family.				
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).				
FT TRANSMEM 58 78 POTENTIAL.				
FT TRANSMEM 122 142 POTENTIAL.				
FT TRANSMEM 191 211 POTENTIAL.				
FT TRANSMEM 218 238 POTENTIAL.				
FT TRANSMEM 299 319 POTENTIAL.				
FT TRANSMEM 328 348 POTENTIAL.				
FT DOMAIN 349 712 CYTOPLASMIC (POTENTIAL).				
FT TRANSMEM 713 733 POTENTIAL.				
FT TRANSMEM 758 778 POTENTIAL.				
FT TRANSMEM 834 854 POTENTIAL.				
FT TRANSMEM 855 875 POTENTIAL.				
FT TRANSMEM 938 958 POTENTIAL.				
FT TRANSMEM 975 995 POTENTIAL.				
FT DOMAIN 996 1281 CYTOPLASMIC (POTENTIAL).				
FT NIP_BIND 429 436 ATP (POTENTIAL).				
FT NP_BIND 1071 1078 ATP (POTENTIAL).				
SQ SEQUENCE 140866 AA; 1281 AA; Length 1281;				
Query Match	Score 99;	DB 1;	Length	1281;

{5} SEQUENCE FROM N.A.
RP TISSUE=Lung;
RC MEDLINE=2382257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Shevchenko C.M., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Buetow K.H., Schaefer C.F., Bhat N.K., Altenschul S.F., Zeeberg B., Buetow K.H., Moore T., Max S.I., Wang J., Hsieh P., Hopkins R.F., Jordan B., Farmer A.A., Rubin G.M., Hong L., Diatchenko L., Marsusina K., Casavant T.L., Scheetz T.E., Stapleton M., Soares M.B., Bonaldo M.F., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Brownstein M.J., Usdin T.B., Toshiyuki S., Carmirico P., Prange C., Bosak S., Raha S., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Huijyk S.W., Vailan D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Palley J., Heitton B., Keittman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., Butterfield Y.S.N., Krywinski M.I., Skalska U., Smailus D.E., Scherich A., Schein J.E., Jones S.J.M., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). [6]

RP SEQUENCE OF 2B-855 FROM N.A.
RC TISSUE=Spleen;
RA Ohara O., Nagase T., Kikuno R., Okumura K.; RT The nucleotide sequence of a long cDNA clone isolated from human spleen.; RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases. [7]

RN SEQUENCE OF 101-855 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039618; PubMed=10574461;
RA Hirosewa M., Nagase T., Ishikara K.-I., Kikuno R., Momura N., RA "Characterization of cDNA clones selected by the GeneMark analysis from size-fractionated cDNA libraries from human brain." RT DNA Res. 6:329-336(1999).
RL -1- FUNCTION: Involved in transcription-coupled repair (TCR) and transcription.
CC -1- SUBUNIT: Associates with RNA polymerase II, the TCR-specific proteins CRN1/CSA and ERCC6/CSB, and XPA.
CC -1- SUBUNITULAR LOCATION: Nuclear (Probable).
CC -1- SIMILARITY: Contains 11 TPR repeats.
CC -1- SIMILARITY: Contains 10 TPR repeats.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to frame-shifts.
CC -1- CAUTION: Ref.6 sequence differs from that shown due to frame-shifts.
CC -1- CAUTION: This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license/agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC DR EMBL; AF26111; BAB15807.1; .
DR EMBL; AF26051; AAP86951.1; ALT_FRAME.
DR EMBL; AF258567; AAC23770.1; .
DR EMBL; BC07204; AAC07208.1; .
DR EMBL; BC08778; AAC08778.1; ALT_INIT.
DR EMBL; AK07403; BAB84851.1; ALT_FRAME.
DR EMBL; AB033003; BAA86491.1; .
DR GO; GO:0005634; C:nucleus; IC.
DR GO; GO:0005515; P:protein binding; IPI.
DR GO; GO:006233; P:transcription-coupled nucleotide-excision r. . . ; IDA.
DR InterPro; IPR003107; HAT.

{5} SEQUENCE FROM N.A.
RP TISSUE=Lung;
RC MEDLINE=12477932; PMID=12477932; HAT: 6;
RX DNA repair; Transcription; Nuclear protein; Repeat; TPR repeat; DR InterPro; IPR003941; TPR-like.
RA SMART; SM00386; HAT: 6;
RA DNA repair; Transcription; Nuclear protein; Repeat; TPR repeat; DR InterPro; IPR004440; TPR.
RA KWW Polymorphism; FT REPEAT 15 47 HAT 1.
RA KWW FT REPEAT 33 67 TPR 1.
RA KWW FT REPEAT 48 80 HAT 2.
RA KWW FT REPEAT 90 122 HAT 3.
RA KWW FT REPEAT 110 143 TPR 2.
RA KWW FT REPEAT 124 158 HAT 4.
RA KWW FT REPEAT 146 179 TPR 3.
RA KWW FT REPEAT 160 192 HAT 5.
RA KWW FT REPEAT 256 289 TPR 4.
RA KWW FT REPEAT 270 305 HAT 6.
RA KWW FT REPEAT 355 388 TPR 5.
RA KWW FT REPEAT 369 407 HAT 7.
RA KWW FT REPEAT 395 428 TPR 6.
RA KWW FT REPEAT 433 466 TPR 7.
RA KWW FT REPEAT 484 517 TPR 8.
RA KWW FT REPEAT 498 530 HAT 9.
RA KWW FT REPEAT 519 551 TPR 9.
RA KWW FT REPEAT 532 566 HAT 9.
RA KWW FT REPEAT 554 590 TPR 10.
RA KWW FT REPEAT 679 713 HAT 11.
RA KWW FT VARIANT 126 126 V -> I.
RA KWW /FTId=VAR_016248.
RA KWW R -> Q.
RA KWW /FTId=VAR_016249.
RA KWW A -> T.
RA KWW /FTId=VAR_016250.
RA KWW FT CONFLICT 68 68 Y -> T (IN REF. 1).
RA KWW FT CONFLICT 140 140 L -> M (IN REF. 1).
RA KWW FT CONFLICT 447 447 E -> K (IN REF. 5; AAH08778).
RA KWW FT CONFLICT 680 680 E -> K (IN REF. 2).
RA KWW FT CONFLICT 751 753 SAT -> IP (IN REF. 2).
SQ SEQUENCE 955 AA; 100009 MW; CP76917CD65F6D CRC64;
Query Match Score 96.5; DB 1; Length 855;
Best Local Similarity 4.1%; Pred. No. 4.7;
Matches 70; Conservative 49; Mismatches 114; Indels 99; Gaps 16;

Qy 6.67 CELKSMGAVRGCCRENNTPPSL-----BAAIIIPPPDDEKPVYDQMEAWFA 215
Db 66 CSYKLKTRYLKARRAQVKHRCVCTDPAEVNNCHERRAFENHK--MFRILWLYCCFLMDQ 123
Qy 216 GKKVATR-FVFGVTAISIODE-----AKSSEVPKPSRVHATVGFMLKHLIAASR 265
Db 124 GRVTHTRPTEDEFLALPITQHSRISIWPFLYRERLIFPE-TAVRGPRFLKLISPEAAE 182
Qy 265 ALTSGTTISTRLSIAQNLTRMMNTEVLLNQALIELSHTTPEISDILKLC 325
Db 183 YIEYKLSKSDRDEAQ--RLATVNDERFVSKA-----GKSNYQWH-----ELC 225
Qy 326 DLVNLNLGSVKGCONTDYFETPKG--KXKXGRM-CXKLDFQRTMSMELAPDLY---
Db 226 DLISQNPDKVQSLNVD--AIRGGLTRFTDQGKLUWCSLADDYIYRSGHFEKARDYEEAI 283
Qy 376 -----LFSSTWTFNFNPFLDGMGRTSWIGVAGKIRASCKETILVETQCGSGIEA 424
Db 284 RTVMTYRDFEQYDQFEESM-----IAAKRNETAS-----ELGREGED 323
Qy 425 WVNLE-----EERKA-----MLEQDPH 441
Db 324 DVDEELRLARFQLISRPLLINSVLLRQDPH 355
Result 13
XAB2_MOUSE 1D
XAB2_MOUSE 1D
Q9DOD2; Q8VDT5; STANDARD; AC
Q9CVD8; PRT; 855 AA.

28-FEB-2003 (Rel. 41; Created)	DR	EMBL; AK008628; BAB25790.1; -
28-FEB-2003 (Rel. 41; Last sequence update)	DR	EMBL; BC021341; AAH21341.1; -
19-OCT-2003 (Rel. 42; Last annotation update)	DR	MGD; MGI:1914689; 0610041014R1X.
XPA2.	DR	GO; GO:0005634; C:nucleus; IC.
Mus musculus (Mouse)	DR	GO; GO:0005515; P:protein binding; ISS.
Bukurota, Metarota; Chordata; Craniata; Vertebrata; Euteleostomi;	DR	GO; GO:0016283; P:transcription-coupled nucleotide-excision r. - ; ISS.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	DR	InterPro; IPR031107; HAT.
NCBI_TaxID:10090; [1]	DR	InterPro; IPR08940; Prexyl trans.
SEQUENCE FROM N.A.	DR	InterPro; IPR08941; TPR-like.
STRAIN=C57BL/6J; TISSUE=Kidney, and Stomach;	DR	InterPro; IPR01440; TPR.
MEDLINE=20085660; PubMed=11217851;	DR	SMART; SM00386; HAT; 10.
Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Hara A., Fukuda H., Kiyosawa H., Kondo S., Yamamoto T., Aizawa K., Izawa M., Nishi K., Kiyohara H., Bono H., Kasukawa T., Saito T., Okazaki Y., Gojobori T., Bono H., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batallow S., Casavant T., Fleischmann R., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L., Staubli F., Suzui R., Tomita M., Wagner B., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blaake J., Boffelli D., Bojunga R., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fleischer C., Fujita M., Garibaldi M., Gustinovich S., Hill D., Hofmann M., Hume D.A., Kanaya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombret P., Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N., Sasai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang X.H., Weitz C., Whitaker C., Wilming L., Wynnshaw-Borit A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;	KW	DNA repair; Transcription; Nuclear protein;
"Functional annotation of a full-length mouse cDNA collection.";	DR	Repeat; TPR repeat.
Nature 409:695-696(2001). [2]	DR	REPEAT 15 . 47 HAT 1.
SEQUENCE FROM N.A.	DR	REPEAT 33 . 67 HAT 2.
MEDLINE=22388257; PubMed=12477932;	DR	REPEAT 48 . 80 HAT 2.
Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Blatchley M., Soares M.B., Bonaldo M.F., Cabantog T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Carninci P., Prange C., Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKerrow J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Bahey J., Heaton R., Kettman M., Madan A., Rodriguez S., Sanchez A., Blakesley R.W., Shewchenko Y., Bouffard G.G., Rodriguez A.C., Grimes J.W., Schmutz J., Myer R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.B., Schneirach A., Schein J.E., Jones S.J.M., Maitra M.A.; *Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. ";	DR	REPEAT 90 . 122 HAT 3.
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	DR	REPEAT 110 . 143 HAT 4.
-!- FUNCTION: Involved in transcription-coupled repair (TCR) and transcription (By similarity).	DR	REPEAT 124 . 158 HAT 5.
-!- SUBCELLULAR LOCATION: Nuclear (Probable).	DR	REPEAT 146 . 179 HAT 3.
-!- SIMILARITY: Contains 11 HAT repeats.	DR	REPEAT 160 . 192 HAT 5.
-!- SIMILARITY: Contains 10 TPR repeats.	DR	REPEAT 256 . 289 HAT 4.
SEQUENCE FROM N.A.	DR	REPEAT 270 . 305 HAT 6.
MEDLINE=20085660; PubMed=11217851;	DR	REPEAT 355 . 388 HAT 5.
Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Hara A., Fukuda H., Kiyosawa H., Kondo S., Yamamoto T., Aizawa K., Izawa M., Nishi K., Kiyohara H., Bono H., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batallow S., Casavant T., Fleischmann R., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L., Staubli F., Suzui R., Tomita M., Wagner B., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blaake J., Boffelli D., Bojunga R., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fleischer C., Fujita M., Garibaldi M., Gustinovich S., Hill D., Hofmann M., Hume D.A., Kanaya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombret P., Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N., Sasai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang X.H., Weitz C., Whitaker C., Wilming L., Wynnshaw-Borit A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;	DR	REPEAT 369 . 407 HAT 7.
"Functional annotation of a full-length mouse cDNA collection.";	DR	REPEAT 395 . 426 HAT 6.
Nature 409:695-696(2001). [2]	DR	REPEAT 433 . 466 HAT 7.
SEQUENCE FROM N.A.	DR	REPEAT 484 . 517 HAT 9.
MEDLINE=22388257; PubMed=12477932;	DR	REPEAT 498 . 530 HAT 8.
Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Blatchley M., Soares M.B., Bonaldo M.F., Cabantog T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Carninci P., Prange C., Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKerrow J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Bahey J., Heaton R., Kettman M., Madan A., Rodriguez S., Sanchez A., Blakesley R.W., Shewchenko Y., Bouffard G.G., Rodriguez A.C., Grimes J.W., Schmutz J., Myer R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.B., Schneirach A., Schein J.E., Jones S.J.M., Maitra M.A.; *Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. ";	DR	REPEAT 519 . 551 HAT 9.
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	DR	REPEAT 532 . 566 HAT 8.
-!- FUNCTION: Involved in transcription-coupled repair (TCR) and transcription (By similarity).	DR	REPEAT 554 . 590 HAT 10.
-!- SUBCELLULAR LOCATION: Nuclear (Probable).	DR	REPEAT 571 . 605 HAT 10.
-!- SIMILARITY: Contains 11 HAT repeats.	DR	REPEAT 679 . 713 HAT 11.
SEQUENCE FROM N.A.	DR	CONFLICT 684 . 684 A -> T (IN REF. 2).
MEDLINE=22388257; PubMed=12477932;	DR	CONFLICT 842 . 842 Q -> L (IN REF. 1; BAB257-0).
Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Hara A., Fukuda H., Kiyosawa H., Kondo S., Yamamoto T., Aizawa K., Izawa M., Nishi K., Kiyohara H., Bono H., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batallow S., Casavant T., Fleischmann R., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L., Staubli F., Suzui R., Tomita M., Wagner B., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blaake J., Boffelli D., Bojunga R., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fleischer C., Fujita M., Garibaldi M., Gustinovich S., Hill D., Hofmann M., Hume D.A., Kanaya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombret P., Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N., Sasai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang X.H., Weitz C., Whitaker C., Wilming L., Wynnshaw-Borit A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;	DR	CONFLICT 855 . 99987 MW; 6A5D6A7474E7FCBD CRC64;
"Functional annotation of a full-length mouse cDNA collection.";	DR	SEQUENCE 70 . Conservative 49 ; Mismatches 114 ; Indels 99 ; Gaps 16 ; Best Local Similarity 21.1% ; Score 96.5 ; DB 1; Length 855;
Nature 409:695-696(2001). [2]	DR	Query Match 167 CIPKNSWGAVERGCRENITHPSLS --- EAALEPPRDLPEKXYDQMPALWFA 215
SEQUENCE FROM N.A.	DR	Db 66 CSYKLWWRYLKARRAQKHRCVTDPAVEDYNNCHCERAFYEMHK-MERLWLDYCFMLMDQ 123
MEDLINE=22388257; PubMed=12477932;	DR	Query Match 216 GKKVATPR-FVEYVKAISIODE --- AKSESVPKPSRVAWTFPLWKLILLASR 265
Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Blatchley M., Soares M.B., Bonaldo M.F., Cabantog T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Carninci P., Prange C., Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKerrow J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Bahey J., Heaton R., Kettman M., Madan A., Rodriguez S., Sanchez A., Blakesley R.W., Shewchenko Y., Bouffard G.G., Rodriguez A.C., Grimes J.W., Schmutz J., Myer R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.B., Schneirach A., Schein J.E., Jones S.J.M., Maitra M.A.; *Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. ";	DR	Db 124 GKVTHTRTFEDRALPITOHSHRIPYLFLRFLSHPE-TAVYGRYRFLKLSPSESAEE 182
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	DR	Query Match 266 ALTSGTISTRSLTAAQVNLRTTMNTEYLDTQNLNTGQNLFWQAQALLESHTPEISDLKLC 325
-!- FUNCTION: Involved in transcription-coupled repair (TCR) and transcription (By similarity).	DR	Db 183 YIEYLKSSDRDDEAAQ-RLATVYNDERFVSKA --- GKSNYQLNE- --- EJK 225
-!- SUBCELLULAR LOCATION: Nuclear (Probable).	DR	Query Match 326 DLVNLINGSVKCQENGDFETPKG --- KEGYGRM-CBFLDFQRTMSSMERADIV --- 375
-!- SIMILARITY: Contains 11 HAT repeats.	DR	Db 226 DLISQNEPDKVQLSNNVDA-ATIRGGLTLKFTDQLGKRLWCSLADYYTRSGFKEKARDVYEEAI 2B3
SEQUENCE FROM N.A.	DR	Query Match 376 --- LFSSWNTNNFNFNLDFFGGRTSMGVAKXIESACKFLILVPTQCGSSGIEA 424
MEDLINE=22388257; PubMed=12477932;	DR	Db 284 RTVMTYTRDFTQDSTQFBEIS- --- IAKMKTAS- --- ELGRBEEED 323
Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Hara A., Fukuda H., Kiyosawa H., Kondo S., Yamamoto T., Aizawa K., Izawa M., Nishi K., Kiyohara H., Bono H., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batallow S., Casavant T., Fleischmann R., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L., Staubli F., Suzui R., Tomita M., Wagner B., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blaake J., Boffelli D., Bojunga R., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fleischer C., Fujita M., Garibaldi M., Gustinovich S., Hill D., Hofmann M., Hume D.A., Kanaya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombret P., Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N., Sasai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang X.H., Weitz C., Whitaker C., Wilming L., Wynnshaw-Borit A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;	DR	Query Match 425 WNLNLS- --- EKOMKA- --- MEBQDPH 441
"Functional annotation of a full-length mouse cDNA collection.";	DR	Db 324 DVDLERLARFEQLISRPLRLLNSVLRQNPHE 355
Nature 409:695-696(2001). [2]	DR	RESULT 14 XAB2 RAT ID XAB2 RAT STANDARD; PRT; 855 AA.
SEQUENCE FROM N.A.	DR	AC Q99P01; 28-FEB-2003 (Rel. 41; Last sequence update)
MEDLINE=20085660; PubMed=11217851;	DR	AC Q99P01; 28-FEB-2003 (Rel. 41; Last annotation update)
XPA2.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Mus musculus (Mouse)	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Bukurota, Metarota; Chordata; Craniata; Vertebrata; Euteleostomi;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
NCBI_TaxID:10090; [1]	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
SEQUENCE FROM N.A.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
STRAIN=C57BL/6J; TISSUE=Kidney, and Stomach;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
MEDLINE=20085660; PubMed=11217851;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Hara A., Fukuda H., Kiyosawa H., Kondo S., Yamamoto T., Aizawa K., Izawa M., Nishi K., Kiyohara H., Bono H., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batallow S., Casavant T., Fleischmann R., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L., Staubli F., Suzui R., Tomita M., Wagner B., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blaake J., Boffelli D., Bojunga R., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fleischer C., Fujita M., Garibaldi M., Gustinovich S., Hill D., Hofmann M., Hume D.A., Kanaya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombret P., Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N., Sasai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang X.H., Weitz C., Whitaker C., Wilming L., Wynnshaw-Borit A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
"Functional annotation of a full-length mouse cDNA collection.";	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Nature 409:695-696(2001). [2]	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
SEQUENCE FROM N.A.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
MEDLINE=22388257; PubMed=12477932;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Blatchley M., Soares M.B., Bonaldo M.F., Cabantog T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Carninci P., Prange C., Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKerrow J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Bahey J., Heaton R., Kettman M., Madan A., Rodriguez S., Sanchez A., Blakesley R.W., Shewchenko Y., Bouffard G.G., Rodriguez A.C., Grimes J.W., Schmutz J., Myer R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.B., Schneirach A., Schein J.E., Jones S.J.M., Maitra M.A.; *Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. ";	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- FUNCTION: Involved in transcription-coupled repair (TCR) and transcription (By similarity).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- SUBCELLULAR LOCATION: Nuclear (Probable).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- SIMILARITY: Contains 11 HAT repeats.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
SEQUENCE FROM N.A.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
MEDLINE=22388257; PubMed=12477932;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Hara A., Fukuda H., Kiyosawa H., Kondo S., Yamamoto T., Aizawa K., Izawa M., Nishi K., Kiyohara H., Bono H., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batallow S., Casavant T., Fleischmann R., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L., Staubli F., Suzui R., Tomita M., Wagner B., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blaake J., Boffelli D., Bojunga R., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fleischer C., Fujita M., Garibaldi M., Gustinovich S., Hill D., Hofmann M., Hume D.A., Kanaya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombret P., Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N., Sasai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang X.H., Weitz C., Whitaker C., Wilming L., Wynnshaw-Borit A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
"Functional annotation of a full-length mouse cDNA collection.";	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Nature 409:695-696(2001). [2]	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
SEQUENCE FROM N.A.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
MEDLINE=22388257; PubMed=12477932;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Blatchley M., Soares M.B., Bonaldo M.F., Cabantog T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Carninci P., Prange C., Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKerrow J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Bahey J., Heaton R., Kettman M., Madan A., Rodriguez S., Sanchez A., Blakesley R.W., Shewchenko Y., Bouffard G.G., Rodriguez A.C., Grimes J.W., Schmutz J., Myer R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.B., Schneirach A., Schein J.E., Jones S.J.M., Maitra M.A.; *Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. ";	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- FUNCTION: Involved in transcription-coupled repair (TCR) and transcription (By similarity).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- SUBCELLULAR LOCATION: Nuclear (Probable).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- SIMILARITY: Contains 11 HAT repeats.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
SEQUENCE FROM N.A.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
MEDLINE=22388257; PubMed=12477932;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Hara A., Fukuda H., Kiyosawa H., Kondo S., Yamamoto T., Aizawa K., Izawa M., Nishi K., Kiyohara H., Bono H., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batallow S., Casavant T., Fleischmann R., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L., Staubli F., Suzui R., Tomita M., Wagner B., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blaake J., Boffelli D., Bojunga R., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fleischer C., Fujita M., Garibaldi M., Gustinovich S., Hill D., Hofmann M., Hume D.A., Kanaya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombret P., Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N., Sasai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang X.H., Weitz C., Whitaker C., Wilming L., Wynnshaw-Borit A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
"Functional annotation of a full-length mouse cDNA collection.";	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Nature 409:695-696(2001). [2]	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
SEQUENCE FROM N.A.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
MEDLINE=22388257; PubMed=12477932;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Blatchley M., Soares M.B., Bonaldo M.F., Cabantog T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Carninci P., Prange C., Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKerrow J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Bahey J., Heaton R., Kettman M., Madan A., Rodriguez S., Sanchez A., Blakesley R.W., Shewchenko Y., Bouffard G.G., Rodriguez A.C., Grimes J.W., Schmutz J., Myer R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.B., Schneirach A., Schein J.E., Jones S.J.M., Maitra M.A.; *Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. ";	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- FUNCTION: Involved in transcription-coupled repair (TCR) and transcription (By similarity).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- SUBCELLULAR LOCATION: Nuclear (Probable).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- SIMILARITY: Contains 11 HAT repeats.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
SEQUENCE FROM N.A.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
MEDLINE=22388257; PubMed=12477932;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Hara A., Fukuda H., Kiyosawa H., Kondo S., Yamamoto T., Aizawa K., Izawa M., Nishi K., Kiyohara H., Bono H., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batallow S., Casavant T., Fleischmann R., Gaasterland T., Gissi C., King B., Kochiwa H., Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J., Schriml L., Staubli F., Suzui R., Tomita M., Wagner B., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G., Blaake J., Boffelli D., Bojunga R., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fleischer C., Fujita M., Garibaldi M., Gustinovich S., Hill D., Hofmann M., Hume D.A., Kanaya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombret P., Nordone P., Ring B., Ringwald M., Rodriguez T., Sakamoto N., Sasai H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyo-oka K., Wang X.H., Weitz C., Whitaker C., Wilming L., Wynnshaw-Borit A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
"Functional annotation of a full-length mouse cDNA collection.";	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Nature 409:695-696(2001). [2]	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
SEQUENCE FROM N.A.	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
MEDLINE=22388257; PubMed=12477932;	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P., Blatchley M., Soares M.B., Bonaldo M.F., Cabantog T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Carninci P., Prange C., Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKerrow J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Bahey J., Heaton R., Kettman M., Madan A., Rodriguez S., Sanchez A., Blakesley R.W., Shewchenko Y., Bouffard G.G., Rodriguez A.C., Grimes J.W., Schmutz J., Myer R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.B., Schneirach A., Schein J.E., Jones S.J.M., Maitra M.A.; *Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. ";	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- FUNCTION: Involved in transcription-coupled repair (TCR) and transcription (By similarity).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- SUBCELLULAR LOCATION: Nuclear (Probable).	DR	AC Q99P01; 28-FEB-2003 (Rel. 42; Last annotation update)
-!- SIMILARITY: Contains		

CC SHOWN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatic Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC	Db	958	N-PEKINLALRMLGILGSVQRLIKEQVGLIAEFCPTFLDSIAGTISKW	1002
CC	Search completed: July 6, 2004, 13:37:16			
CC	Job time : 13.6195 secs			
CC				
CC	DR	14114; AAA45475.1; -.		
CC	DR	AAA45476.1; -.		
CC	DR	AAA45477; AAA45465.1; -.		
CC	DR	EMBL; M14707; AAA45466.1; ALT_INIT.		
CC	DR	EMBL; M16632; AAA45471.1; -.		
CC	DR	PIR; A03305; A03905.		
CC	DR	PIR; A25941; GNNYEM.		
CC	DR	PIR; A94149; GNNYEM.		
CC	DR	PDB; 1HAY; 23-DEC-96.		
CC	DR	MBiOPS; C03.005; -.		
CC	DR	InterPro; IPR004004; Calci_poli_biol.		
CC	DR	InterPro; IPR009003; Cys_Ser_trypsin.		
CC	DR	InterPro; IPR006065; RNA_helicase.		
CC	DR	InterPro; IPR007095; RNA_pol_DS_PS.		
CC	DR	InterPro; IPR001205; RNA_pol_P3D.		
CC	DR	InterPro; IPR007094; RNA_pol_Psvr.		
CC	DR	InterPro; IPR008975; Viral_cap_coat.		
CC	DR	Pfam; PF00680; RNA_dep_RNA_pol; 1.		
CC	DR	Pfam; PF00910; RNA_helicase; 1.		
CC	DR	PRINTS; PR00918; CALICIVIRUSNS.		
CC	KW	Polyprotein; Coat_protein; Core_protein; Transfase;		
CC	KW	RNA-directed RNA_polymerase; Hydrolase; Thiol_protease; 3D-structure.		
CC	FT	CHAIN 1 23 COAT PROTEIN VP4 (P1A).		
CC	FT	CHAIN 24 245 COAT PROTEIN VP2 (P1B).		
CC	FT	CHAIN 246 491 COAT PROTEIN VP3 (P1C).		
CC	FT	CHAIN 492 836 COAT PROTEIN VP1 (P1D).		
CC	FT	CHAIN 837 980 CORE PROTEIN P2A.		
CC	FT	CHAIN 981 1087 CORE PROTEIN P2B.		
CC	FT	CHAIN 1088 1422 CORE PROTEIN P2C.		
CC	FT	CHAIN 1423 1496 PROBABLE PROTEIN P3A.		
CC	FT	CHAIN 1497 1519 PROBABLE PROTEIN P3B.		
CC	FT	CHAIN 1520 1738 PROBABLE PROTEIN P3C.		
CC	FT	CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.		
CC	FT	VARIANT 77 77 K -> R (IN ATTENUATED STRAIN).		
CC	FT	VARIANT 764 764 B -> V (IN ATTENUATED STRAIN).		
CC	FT	VARIANT 821 821 N -> S (IN ATTENUATED STRAIN).		
CC	FT	VARIANT 1052 1052 A -> V (IN ATTENUATED STRAIN).		
CC	FT	VARIANT 1062 1062 G -> A (IN ATTENUATED STRAIN).		
CC	FT	VARIANT 1118 1118 K -> M (IN ATTENUATED STRAIN).		
CC	FT	VARIANT 1151 1151 E -> K (IN ATTENUATED STRAIN).		
CC	FT	VARIANT 1163 1163 F -> S (IN ATTENUATED STRAIN).		
CC	FT	VARIANT 1277 1277 V -> I (IN ATTENUATED STRAIN).		
CC	FT	VARIANT 1500 1500 H -> Y (IN ATTENUATED STRAIN).		
CC	FT	VARIANT 1805 1805 D -> N (IN ATTENUATED STRAIN).		
CC	FT	VARIANT 1930 1930 S -> T (IN ATTENUATED STRAIN).		
CC	SQ	SEQUENCE 2227 AA; 251506 MW; 01E225EP/AB740A6 CRC64;		
CC	Query Match	4.0% Score 95.5; DB 1; Length 2227;		
CC	Matches 50; Conservative 39; Mismatches 77; Indels 61; Gaps 12;			
Qy	197	PRDDLPKXYVQDMEALMPAGKCYATRPFVEGKAISSTQDEAKSSEVPKESTV-----	249	
Db	782	PRSEBEDRFESIE---CRKPYKELEVKORYKAQETSNKEUTPPPRMKGLFSQA 937		
Qy	250	-----HAYTGFLWHLIILASA-----LTSGCTSTLTLIAQAYNLR-TRNNME 292		
Db	838	KISLSPYTERHEMKFSRGTIDTTRREGFLAAGRSVWTLLEMAGVLTCLIRINDE 897		
Qy	293	-----TYLDNATGFLFW----WQAILEL-----SHTTPRISDLKLICDILWLL 331		
Db	898	KTTEMDDKDV1VLLKEFISNKWVSKVNF2HGMELTIAANSKDFPMSETOLCPJLHML 957		
Qy	332	NGSVKQCN-GDYFETFKG---KEGYGRMCE -- YLD-FORTMSSM 368		

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 13:33:50 ; Search time 13.8278 Seconds

Title: US-09-857-518A-31
Perfect score: 1758

Sequence: ETGATDVREKTYC3VHSD.....ERVVKDVRPRFVIDVENTL 326

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 9619526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : PIR_78;*

1: pir1;*

2: pir2;*

3: pir3;*

4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	971	55.2	361	2 T12571	cinnamyl-alcohol d
2	962	54.8	337	2 T05625	cinnamyl-alcohol d
3	959	54.6	359	1 S28043	cinnamyl-alcohol d
5	881	53.2	354	2 S72477	probable cinnamyl-alcohol d
6	881	50.1	356	2 S71179	cinnamyl-alcohol d
7	881	50.1	360	2 T08581	cinnamyl-alcohol d
8	853	48.5	363	2 T05624	cinnamyl-alcohol d
9	819.5	46.6	375	2 D84605	alcohol dehydrogen
10	800	45.6	376	2 B84604	cinnamyl-alcohol d
11	740	42.1	357	2 T05413	cinnamyl-alcohol d
12	739	42.0	357	2 S49443	cinnamyl-alcohol d
13	719	40.9	357	2 S49444	cinnamyl-alcohol d
14	686	39.0	357	2 S23526	cinnamyl-alcohol d
15	682	38.7	357	2 S23525	cinnamyl-alcohol d
16	666	37.9	357	2 T09141	cinnamyl-alcohol d
17	645	36.7	357	2 T05413	cinnamyl-alcohol d
18	641	36.5	358	2 S31572	cinnamyl-alcohol d
19	629	35.8	354	2 S60424	cinnamyl-alcohol d
20	628	35.7	367	2 T02930	cinnamyl-alcohol d
21	621	35.3	367	2 T02767	cinnamyl-alcohol d
22	616	35.1	362	2 S45094	cinnamyl-alcohol d
23	613.5	34.9	362	2 D87125	alcohol dehydrogen
24	609	34.6	355	2 AC2768	alcohol dehydrogen
25	609	34.6	368	2 E97448	cinnamyl-alcohol d
26	602	34.2	341	2 S31571	alcohol dehydrogen
27	599.5	34.1	346	2 H70860	alcohol dehydrogen
28	599.5	34.1	346	2 JC1376	alcohol dehydrogen
29	596.5	33.9	349	2 C69583	alcohol dehydrogen

RESULT 2

S28045

cinnamyl-alcohol dehydrogenase (BC 1.1.1.195) EL13 - parsley (fragment)

C;Species: Petroselinum crispum

(parsley)

alcohol dehydrogen	30	587.5	33.4	355	2 E96751
alcohol dehydrogen	31	581.5	33.1	353	2 D83361
alcohol dehydrogen	32	571.5	32.5	352	2 G82719
alcohol dehydrogen	33	568.5	32.3	349	2 D82563
alcohol dehydrogen	34	566.5	32.2	348	2 AF2824
alcohol dehydrogen	35	561	31.9	348	2 G82645
probable alcohol d	36	555.5	31.6	349	2 E64759
alcohol dehydrogen	37	551.5	31.4	349	2 C906756
alcohol dehydrogen	38	551.5	31.4	349	2 P8526
alcohol dehydrogen	39	548.5	31.2	349	2 H82643
probable cinnamyl-	40	504	28.7	348	1 H64657
zinc-dependent alc	41	503	28.6	350	2 A71857
probable alcohol d	42	475	27.0	358	2 BB13C2
zinc-dependent alc	43	474	23.6	365	2 H718C8
probable aryl alc	44	415.5	23.6	336	1 S76928
hypothetical prote	45	405.5	23.1	241	2 G97602

ALIGNMENTS

RESULT 1

T12571

Common ice plant

C;Species: Mesembryanthemum crystallinum (common ice plant)

C;Date: 23-Jul-1999 #sequence_change 08-Dec-2000

C;Accession: T12571

R;Michałowski, C.B.; Bohmert, H.J.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z17331

A;Accession: T12571

A;Status: Preliminary; translated from GB/EMBL/DBDJ

A;Molecule Type: mRNA

A;Residues: 1-361 <MIC>

C;Cross-references: EMBL:U79770; PID:G1724110; PID:G1724110

C;Keywords: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology

P;32-340/Domain: Long-chain alcohol dehydrogenase; zinc

F;51,73,167/Binding site: zinc, catalytic (Cys, His, Cys); #sites predicted

Query Match

Best Local Similarity

Matches

190;

Conservative

52;

Mismatches

63;

Indels

38;

Gaps

6;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

;

Db	146	VRIPENPLPLEQAAPLICAGTVFSPMKHFAMTPGKCKGIGLGGVGHKGKIAKF--	203	RESULT 14
				S23526
Qy	165	RPPPLNPLGIGWSRQLSIVPLIKEGGSYSTSPALMHSIIRTDDOMERAMSTMDGIIITV	224	cinnamyl-alcohol dehydrogenase (BC 1.1.1.115) CAD19 - common tobacco
				C;Species: Nicotiana tabacum (common tobacco)
Db	204	-G-HVTVTSSKKEAVMEVGDAY-----LVSQDTKMMEAESLDYIMDTI	252	C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Oct-1999
				C;Accession: S23526
Qy	225	PAVPPLPLISLLKTKNGKNTGAVQPLDPVPLLIKGKMYAGSAAGGMKETQBMIDF	284	R;Knight, M.E.; Halpin, C.; Schuch, W.
				Plant Mol. Biol. 19, 793-801, 1992
Db	253	PVAPPLEPLVLLAKTNGKLVLMGTVPEPLHFTPLIIGRRTAGSPFGGMETQETLDP	312	A;Title: Identification and characterisation of cDNA clones encoding cinnamyl alcohol de
				A;Reference number: S23525; MUID:92353388; PMID:1643282
Qy	285	AAEINITADIEVPIPDYINTAMERVVKDVRFRFIDV	322	A;Accession: S23526
				A;Status: preliminary
Db	313	CAEKVSSMIEVWGLDYINTAMERLEKNDVWVRFVVDV	350	A;Molecule type: mRNA
				A;Residues: 1-357 <KNN>
Qy	349444	cinnamyl-alcohol dehydrogenase (BC 1.1.1.195) B - loblolly pine	349444	A;Cross-references: EMBL:XC62344; NID:919840; PIDN:CAA44217.1; PID:919841
				C;Genetics:
Db	A;Gene: CAD19	C;Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology	A;Keywords: alcohol metabolism; oxidoreductase; zinc	
				F;32-340/Domain: long-chain alcohol dehydrogenase homology <LADH>
Qy	55285	#sequence_revision 20-Feb-1999 #text_change 15-Oct-1999	F;47,69,163/Binding site: zinc, catalytic (Cys, His, Cys) #status predicted	R;Knight, J.J.; Liu, W.; Whetten, R.; Sederoff, R.; O'Malley, D.
				submitted to the EMBL Data Library, September 1994
A;Description: Genetic analysis of cad in loblolly pine. Single gene inheritance, molecular	A;Reference number: S49443	A;Accession: S49444	A;Molecule type: mRNA	A;Introns: 1-357 <MAC>
				A;Cross-references: EMBL:Z37992; NID:9558386; PIDN:CAA86073.1; PID:9558387
Db	R;Knight, J.J.; Liu, W.; Whetten, R.; Sederoff, R.R.; O'Malley, D.M.	A;Keywords: alcohol metabolism; oxidoreductase; zinc	A;Accession: S55285; MUID:95327049; PMID:7603432	Mol. Cell. Genet. 247, 537-545, 1995
				A;Title: Genetic analysis of cinnamyl alcohol dehydrogenase in loblolly pine: single gene
Qy	89	DTGVGULLVJUGSCRNCGPCKREIQQCNCKIMNCNDVYTGDGKPTQGGFANSMVDQNFYK	148	A;Accession: S55286
				A;Molecule type: mRNA
Db	119	IPONPLPLGAAPLICAGTTSPDWWYGGDLPKPMELGAEWPRFRSRRPLNDPGLWGSR	178	A;Cross-references: EMBL:Z37992
				C;Genetics:
Qy	149	IPGGMAPPDAAAPLICAGTTSPDNHGFQFNQSFPRGGI-----LGCGVGHNKCV	197	A;Introns: 123/3; 198/3
				C;Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
Db	198	K-----IKAMGHEHVTVISSSNKQRQELEHLGADDVYSSDTQRAADSDDYDITDV	252	C;Keywords: alcohol metabolism; metalloprotein; oxidoreductase; zinc
				F;32-340/Domain: long-chain alcohol dehydrogenase homology <LADH>
Qy	225	PAVRPLEPLISLKTNGKTVTGVIAQFDLDPVPLI-IGKROMYAGSAIGGMKETQEMID	283	A;Accession: S47,69,163/Binding site: zinc, catalytic (Cys, His, Cys) #status predicted
				A;Introns: 40-9%; Score: 719; DB 2; Length 357;
Db	253	PVCHEPLIYLSLKLIDKGKLLIGVINTLQF-ISPMLWGRKSITGTSIGMSKETEEMID	311	A;Description: Best Local Similarity 41.6%; Pred. No. 3e-52; Mismatches 63; Indels 5; Matches 143; Conservative
				Qy 3 GATDVRKVLVYCVHSIDHARNDWGTSTYPVPGHBLGVVTVGCKVKGKPSWRQ-- 60
Db	33	GVEDVTKVYICGICDPLWQKRNEMKSHPVPVGFHVGTVTE18SEVRGPKVGHVG	92	Db 33 GEDVTKVYICGICDPLWQKRNEMKSHPVPVGFHVGTVTE18SEVRGPKVGHVG
				Qy 61 ----GRCMLHGRRLRPTCENCITHLHENYCPNLQIQTGSKYVNTMGGYSNNMVTDBHF
Db	93	WGTIVGSC-----RSGCAGNQSMBGYCSKRLWTDVNHGTPGFGFASMVQDQMF	115	115 145
				Db 93 WGTIVGSC-----RSGCAGNQSMBGYCSKRLWTDVNHGTPGFGFASMVQDQMF
Qy	116	YRIPONPLPLGAAPLICAGTTSPDWWYGGDLPKPMELGAEWPRFRSRRPLNDPGLWG	175	Qy 116 YRIPONPLPLGAAPLICAGTTSPDWWYGGDLPKPMELGAEWPRFRSRRPLNDPGLWG
				Db 146 YRIPENPLPLEQAPLICAGTTSPMVKHFASTEPEGRKCGI-----LGCGVGH
Qy	176	SRLQSLVPLPLGAAPLICAGTTSPDWWYGGDLPKPMELGAEWPRFRSRRPLNDPGLWG	194	Qy 176 SRLQSLVPLPLGAAPLICAGTTSPDWWYGGDLPKPMELGAEWPRFRSRRPLNDPGLWG
				Db 247 YMDTIPVAPHEPEPLVALLKINGKLVMLGVVDEPLHVTPLJLGRSIAFSGMSNET
Qy	219	GLIDTYPBAVRPLEPLISLKTNGKTVTGVIAQFDLDPVPLI-IGKROMYAGSAIGGMKET	278	Qy 219 GLIDTYPBAVRPLEPLISLKTNGKTVTGVIAQFDLDPVPLI-IGKROMYAGSAIGGMKET
				Db 307 QETLDFCAEKCIVSSMTEVVGJLQYDINTAMKRLKEKNDVYRFVVDV
Qy	247	YMDTIPVAPHEPEPLVALLKINGKLVMLGVVDEPLHVTPLJLGRSIAFSGMSNET	350	Qy 247 YMDTIPVAPHEPEPLVALLKINGKLVMLGVVDEPLHVTPLJLGRSIAFSGMSNET
				Db 307 QETLDFCAEKCIVSSMTEVVGJLQYDINTAMKRLKEKNDVYRFVVDV
Db	312	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 312 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Qy	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV	350	Db 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
				Qy 350 PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVDV
Db	350	PCKEGKVYQSIEVKMYINTAMERLEKNDVYRFVVD		

Query Match 38.7%; Score 681; DB 2; Length 357;
 Best Local Similarity 43.1%; Pred. No. 4-5e-49;
 Matches 146; Conservative 54; Mismatches 101; Indels 38; Gaps 8;

Qy 2 TGATDVRFXLYCCVCHSDIHMKAHDWGTSPTVPGHFLGVYTHEVGCKVKKFKSWROS 61
 Db 32 TGPEDVEVKLYCGLIHTDQVQVNDLGMSNYPVPGHFLGVYGEVVBGPVPSERK--VG 88
 Qy 62 RCWLHGRILPTCEN--CTHLENYCPNULJOTYSKYYDQMTTYGGSKMMADTDEHETR 118
 Db 89 DTVEVGLLVSCCRNGPCKRDIEQVCNRKWLNNCNDVYTQCKPTQGGFARSKMUVQKEVVK 148
 Qy 119 IPDMLPLDAGAPLICAGITTSPPRYGLDKPGNHLGVEPVRRFRSRSPNPLPGWSRL 178
 Db 149 IPEKAKPGRQAPLJCAITRYSPLNHFGRQSGURGI-----LGHGAGHMGV 197
 Qy 179 QSLVPEPLIKEKGSSGT-----SPALMH-----SLRRTDQOMEAAMSTMGIIDTV 224
 Db 198 K-----TAKAMGHHTVISSSNKKRQBALELGAADDYLVSSDTDKQEQSDLSLXIIIDTV 252
 Qy 225 PAYRPLEPLISLILXKNGKTYVGLAVQPLLPVFPFL-IGRKMYAGSAGIGGMKETQEMID 283
 Db 253 PVGHPLEPVSLKIDGKLLMGTLINTPLQF-ISMVMDLGRKSITGSFTGSMKETEEMID 311
 Qy 284 FAAFFNITADIEVTPIDYLNNTAMERRVVKCDVRFERFVIDV 322
 Db 312 FCKERKGVTISQIEITKMDYNTAMERLEQNDVRYRFVVDV 350

Search completed: July 6, 2004, 13:39:31
 Job time : 14.8278 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	2373	100.0	452	10 Q9FVFL	Q9FVFL fragaria an
2	2043	86.1	455	10 Q8GM5	Q8GM5 fragaria ve
3	538	22.7	455	10 Q9JU88	Q9JU88 arabidopsis
4	538	22.7	456	10 Q8GW04	Q8GW04 capsicum ch
5	506	21.3	426	10 Q9FV4	Q9FV4 arabidopsis
6	501	21.1	443	10 Q9F140	Q9F140 arabidopsis
7	498	21.0	443	10 Q9U771	Q9U771 arabidopsis
8	498	20.4	446	10 Q9F1M1	Q9F1M1 arabidopsis
9	485	19.7	435	10 Q9F1M4	Q9F1M0 arabidopsis
10	468	19.7	428	10 Q9F1W4	Q9F1W4 arabidopsis
11	454.5	19.2	435	10 Q9F1M4	Q9F1M0 arabidopsis
12	450	19.0	99	10 Q9F1W4	Q9F1W4 arabidopsis
13	448	18.9	435	10 Q9F1M4	Q9F1M0 arabidopsis
14	438	18.5	474	10 Q9F1M4	Q9F1M0 arabidopsis
15	406	17.1	443	10 Q9F1M4	Q9F1M0 arabidopsis
16	404.5	17.0	433	10 Q9F1M4	Q9F1M0 arabidopsis

Qy	4 LEVSINSKHTKPKSTSSTPLQP--YKULLDOLTPAYVPTIVFFYPTIDHIDENLPLQTLAD	Oy	123 KPFMSMBAISDERYPLIGVQVNPF-DSSGIAIGVSYSHKLIDGTTADCFLKSWGAVERGCRC 181
Db	3 LEITVTSQELVKPSPRMLNHPCHHILSFLDOLAPPFLPMPFLFFY---IN---KTNLSD 55	Db	114 -VSPDHADPTEPLKFLVLFKLFKDFKCFAVAVSVSHKICDASLSTPVCSWTAASKGYAD 171
Qy	62 -----LROALSETLTLYPLSGRVKNN--LYIDDFEEGVPYLEARVNQNCMTDFLRLRK 112	Oy	182 NTHPSLSEAALLFPPRD---DLPKTYDOMEALWAECKVATRFFYGVKAISSEDEA 238
Db	56 KERSDEHKISSSEILNLYPLAGRINKSGDVVCNDV--GVSFEAKADCMNSQILENP 113	Db	172 T-VDPFET-VGADYFPADISIEFPPLLYHE----TSKSTKGRFPGSMLIEKLNRA 221
Qy	113 LECLNFVPIKPKPSMBAISDERYPLIGVQVNPF-SGAIIGVSYSHKLIDGTTADCFLKs 171	Oy	239 KS-ESVPKPSRVAHTGFLKELIAASRALSTGTTSTRSLSTIAAQAVNLRTKMNMTETLDN 297
Db	1144 -PNPLNKLHFFHEVSD-VPLT-VOLTFECGLALGIGLSKRLDGLLFVNS 167	Db	222 SSKRVRVQATRESITALLRCLNTKASHSKSGKVKEAAT--QTMDJDRPRVSSLJLHK 278
Qy	172 WGAIVFRCRENTIHPSLSEAALLFPPRDLEPKYDOME-AIWFAGRKKVATRFFYGVKA 230	Oy	298 ATGNLFWQQAILELSEHTPEISDL-KLCDLVNLINGSVKQNCQDVFETEKGKEG---- 351
Db	168 WAAFAEQTQDDEITTPSF-DIAKMFPPCD----TEINLNMTAGTGTENIVTRFVTLRS 220	Db	279 ATGN-FPPFLKKESSRSKMEETBTNSLQTKQELNELIRNDSEDAKSVREAKRASAA 337
Qy	231 ISSIQDEAKSSSVPKPSRVAHTGFLKHLAASRALTSGTTSTRSLIAAQAVNLRTMN 290	Oy	352 -YGRMCYELDFORTMSMSEPAFDIYPLFESSWTF-FNPLDFGMRGTSNTVGAKIESACK 409
Db	221 VESLRRFEGNKKIRATRVEVLISVNSRFRAST--NHDDKTKGYLTHPVNLBRQAD 277	Db	338 MLLSICG-----ISPEKETYAASSWCRMSFYEANFGNGKPVNVAP---DSVDKTT 383
Qy	291 METVLDNATGHNLFWWQAQILLESHTPEI----SDLRLKCDLVNLINGSVKQNCQDVFETF 346	Oy	410 FILVPTOCGSGIETAWVNLEEBXKAMLEQDOPFLALASPKTLI 452
Db	278 ED-1PDRMFGN-----IMRSVTVYMMINENDEKASLVDQMBEIRKLDATVTRKL 329	Db	384 QVVLMDSKDSEG7EARYVTLPEIDMAKYEHDSELVNTATPSPSTI 426
RESULT 12			
Qy	347 KGKEGYGRMCYELDFORTMSMSEPAFDIYPLFESSWTFNPL--DROWGRSTWIGVAGKI 403	TD	023943 PRELIMINARY; PRT; 99 AA.
Db	330 Q-EDNRQG---ELEFLNKVQAISFVNGHIVSVSFTSLICKFPTVYEAFCWKGKPLWVASA---	AC	023943; PRT; 99 AA.
Qy	404 EASACKFILVWPTQCGSGIEMAVNLLEBKAMLEQDOPFLALAS--PXTL 452	DT	01-JAN-1998 (TRIMBLrel. 05; Created)
Db	382 RMSYKVLVAFFIDTKEGDIEMAWNLNDQMSRFEADEBELLRTWSSNPSVM 432	DT	01-JUN-1998 (TRIMBLrel. 05; Last sequence update)
RESULT 13			
Q9FLW4	PRELIMINARY; PRT; 428 AA.	TD	023943; PRT; 99 AA.
AC	Q9FLW4;	AC	023943; PRT; 99 AA.
DT	01-MAR-2001 (TRIMBLrel. 16; Created)	DT	023943; PRT; 99 AA.
DT	01-JUN-2001 (TRIMBLrel. 17; Last annotation update)	DT	01-JUN-1998 (TRIMBLrel. 05; Last sequence update)
DB	Acetyl-CoA:benzylalcohol acetyltransferase-like protein.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
OS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
OC	euroids II; Brassicales; Brassicaceae; Arabidopsis; Arabidopsis.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
OX	NCBI_TaxID=3702;	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
RN		RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
RP	SEQUENCE FROM N.A.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
RC	STRAIN=Columbia;	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
RX	MEDLINE=98890546; PubMed=9628582;	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
RX	SEQUENCE=428 AA;	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
RI	EMBL; AB008720; 1.;	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	Pfam; PF02458; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,
DR	InterPro; IPR003480; Transferase; 1.	RA	NAME; Tichti L., Lepierier M., Marty I., Lelievre J. M.,</

Result No.	Score	Query Match	Length	DB ID	Description
1	1758	100.0	326	3 AAB16445	Aab36445 Strawberry
2	1758	100.0	326	3 AAY79662	Aay79662 Strawberry
3	1007	57.3	355	6 ADA18387	Ada38387 Soybean c.
4	991	56.4	333	3 AAB16444	Aab36444 Strawberry
5	979	55.7	362	5 AAD79807	Aau79807 Sinapyl al.
6	979	55.7	362	5 AAUT79807	Aau79807 Sinapyl al.
7	979	55.7	362	7 ADD39896	Add39896 Quaking a.
8	960	54.6	357	3 AAG29470	Aag29470 Arabidops.
9	960	54.6	357	5 ABB93285	Abb93285 Herbicida
10	959	54.6	359	5 ABB93286	Abb93286 Herbicida
11	881	50.1	360	3 AAG0179	Aag0179 Arabidops.
12	881	50.1	360	5 ABB93203	Abb93203 Herbicida
13	876	49.8	360	3 AAG25506	Aag25506 Arabidops.
14	871	49.5	306	3 AAG29471	Aag29471 Arabidops.
15	869	49.4	360	6 ADA8379	Ada8379 Soybean c.
16	832	47.3	371	6 ADA8389	Ada8389 Wheat cinn.
17	827	47.0	283	3 AAB16447	Aab36447 Strawberry
18	827	47.0	283	3 AAY79664	Aay79664 Strawberry
19	824	46.9	278	3 AAB16446	Aab36446 Strawberry
20	824	46.9	278	3 AAY79663	Aay79663 Strawberry
21	824.5	46.9	278	3 AAY79663	Aay79663 Strawberry
22	819.5	46.6	375	3 AAG32129	Aag32129 Arabidops.
23	819.5	46.6	375	5 ABB91832	Abb91832 Herbicida
24	802	45.6	376	3 AAG2132	Aag32132 Arabidops.
25	800	45.5	361	6 ADA18383	Ada18383 Cctr. cinn.

CC biosynthetic pathway for aliphatic and/or aromatic ester production in
 CC fruit. The nucleotide sequences can be inserted into the genome of a
 CC fruit-producing plant to regulate aliphatic and/or aromatic ester
 CC formation. Aromatic and/or aliphatic esters in microorganisms, plant
 CC cells or plants are produced by inserting thiolase, alcohol acyl
 CC transferase, alcohol dehydrogenase, pyruvate decarboxylase,
 CC aminotransferase and esterase nucleotide sequences into the genome and
 CC feeding the microorganism or plant with alcohol, aldehydes, alpha-keto
 CC acids, or amino acids and fatty acids, and acyl-CoA. The nucleotides and
 CC their proteins can be used in the processed food industry as food
 CC additives to enhance the flavour of syrups, ice-creams, frozen desserts,
 CC yoghurts and confectionery. They are used: as flavouring agents for oral
 CC medications and vitamins; provide flavour and aroma in beverages,
 CC including alcohol; enhance or reduce fruit flavour, aroma, fragrance or
 CC scent; enhance the flavour or aroma of natural, synthetic or artificial flavour
 CC products; for the production of novel combinations of artificial flavour
 CC substances; as antibacterial or anti-fungal agents; as fragrance or
 CC conditioners, leavening agents and fixatives in perfumes, suspension
 CC aids for aluminium salts in anti-perspirant pharmaceuticals, cleaning
 CC products, personal care products and animal care products; as
 CC disinfectant additives; as decreasing solvents for electronics; as insect
 CC pheromones; and as dye carriers, solvents, insect repellents, miticides,
 CC scabicides, plasticisers and deodorants. The present sequence represents
 CC the specifically claimed strawberry alcohol dehydrogenase
 XX

Sequence 326 AA:

Query Match 100.0%; Score 1758; DB 3; Length 326;
 Best Local Similarity 100.0%; Pred. No. 4e-171;
 Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 EIGATIDVRYKVLVCGVCHDIAMKNDWGTSVTPVPGHELVWTVTEVCKVKFKFSWQ 60
 1 ETGATIDVRYKVLVCGVCHDIAMKNDWGTSVTPVPGHELVWTVTEVCKVKFKFSWQ 60
 Db 61 GRCWLHGRDRPTCNCIHLHENYCPNLQYGSKYDGTMTYGSYSNNNNVTDHEHFIYRIP 120
 61 GRCWLHGRDRPTCNCIHLHENYCPNLQYGSKYDGTMTYGSYSNNNNVTDHEHFIYRIP 120
 Qy 121 DNLPLDGAAPLCAAGITTSVSPWRYGLDKPGMHLGVENPRFRSRPPLNPLGLNGSRLQS 180
 Db 121 DNLPLDGAAPLCAAGITTSVSPWRYGLDKPGMHLGVENPRFRSRPPLNPLGLNGSRLQS 180
 Qy 181 IWPPLKIGGSYCTSPALMHSLLRTDQMCMAAMSTMDGIDTPAVPLEPLISLTKN 240
 181 IWPPLKIGGSYCTSPALMHSLLRTDQMCMAAMSTMDGIDTPAVPLEPLISLTKN 240
 Db 241 GKVUTGIAVQDLPVPLIIGRKMVAGSAIGMKETOBMDAEEINITADIEVTPID 300
 241 GKVUTGIAVQDLPVPLIIGRKMVAGSAIGMKETOBMDAEEINITADIEVTPID 300
 Db 301 YINTAMERVVKDVRFRFVIDVENTL 326
 Qy 301 YINTAMERVVKDVRFRFVIDVENTL 326
 Db 301 YINTAMERVVKDVRFRFVIDVENTL 326

RESULT 2

AAV79662 Standard; protein: 326 AA.
 XX

XX AAV79662;
 XX DT 12-SEP-2003 (revised)
 XX DT 29-AUG-2000 (first entry)

XX Strawberry alcohol dehydrogenase SLF193 (C-terminal sequence).
 KW Strawberry; alcohol dehydrogenase; fruit; ripening; ester; flavour;
 KW aroma; transgenic plant.
 XX Fragaria x ananassa.

PN BP1005190-A1.
 XX 07-JUN-2000.
 PD 02-DEC-1998;
 XX 98EP-00204018.
 PF 02-DEC-1998;
 XX 98EP-00204018.
 PR PA
 XX (CPR)- CPRO-DLO CENT PLANTENVEREDELINGS REPROD.
 PI Verhoeven HA, Van Tunen AJ, Aharoni A, Luecker J, O'connell AP;
 XX WPI: 2000-378264/33.
 DR N-PSDB; AAA27672.
 XX New polynucleotides encoding enzymes from the biosynthetic pathway for
 PT aromatic and/or aliphatic ester production in fruit used to modify plant
 PT flavors.
 XX
 PS Claim 36; Page 81-82; 116PP; English.
 XX
 CC The present sequence is that of the C-terminal region of strawberry cv.
 CC Bisantia alcohol dehydrogenase SLF193, an enzyme that shows upregulated
 CC expression in ripening fruit. The invention relates to DNA sequences (see
 CC AYA2766-78) encoding enzymes (see AYA27656-68) involved in the metabolic
 CC pathway leading to the formation of aliphatic and/or aromatic esters in
 CC ripening fruit. The enzymes have alcohol acyl transferase, alcohol
 CC dehydrogenase, pyruvate decarboxylase, thiolase or aminotransferase
 CC activity. Expression vectors comprising the DNA sequences may be used to
 CC regulate ester formation in fruit. Genetically modified plants, plant
 CC cells and microorganisms can be used to produce esters. The DNA
 CC sequences, polypeptides and antibodies are also used to screen fruit for
 CC volatile ester compounds; for quality such as flavour, fragrance, aroma,
 CC scent, texture or shape; to distinguish between cultivars and varieties;
 CC and to monitor harvest time, post-harvest quality, shelf-life, timing of
 CC pesticide application, and resistance capacity based on volatile ester
 CC profiles. (Updated on 12-SEP-2003 to standardise CS Field)

Sequence 326 AA;

Query Match 100.0%; Score 1758; DB 3; Length 326;
 Best Local Similarity 100.0%; Pred. No. 4e-171; Mismatches 0; Indels 0; Gaps 0;
 Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ETGATIDVRYKVLVCGVCHDIAMKNDWGTSVTPVPGHELVWTVTEVCKVKFKFSWQ 60
 1 ETGATIDVRYKVLVCGVCHDIAMKNDWGTSVTPVPGHELVWTVTEVCKVKFKFSWQ 60
 Db 61 GRCWLHGRDRPTCNCIHLHENYCPNLQYGSKYDGTMTYGSYSNNNNVTDHEHFIYRIP 120
 61 GRCWLHGRDRPTCNCIHLHENYCPNLQYGSKYDGTMTYGSYSNNNNVTDHEHFIYRIP 120
 Qy 61 GRCLHGRDRPTCNCIHLHENYCPNLQYGSKYDGTMTYGSYSNNNNVTDHEHFIYRIP 120
 61 GRCLHGRDRPTCNCIHLHENYCPNLQYGSKYDGTMTYGSYSNNNNVTDHEHFIYRIP 120
 Db 61 GKVUTGIAVQDLPVPLIIGRKMVAGSAIGMKETOBMDAEEINITADIEVTPID 300
 61 GKVUTGIAVQDLPVPLIIGRKMVAGSAIGMKETOBMDAEEINITADIEVTPID 300
 Qy 121 DNLPLDGAAPLCAAGITTSVSPWRYGLDKPGMHLGVENPRFRSRPPLNPLGLNGSRLQS 180
 121 DNLPLDGAAPLCAAGITTSVSPWRYGLDKPGMHLGVENPRFRSRPPLNPLGLNGSRLQS 180
 Db 121 DNLPLDGAAPLCAAGITTSVSPWRYGLDKPGMHLGVENPRFRSRPPLNPLGLNGSRLQS 180
 121 DNLPLDGAAPLCAAGITTSVSPWRYGLDKPGMHLGVENPRFRSRPPLNPLGLNGSRLQS 180
 Qy 181 IWPPLKIGGSYCTSPALMHSLLRTDQMCMAAMSTMDGIDTPAVPLEPLISLTKN 240
 181 IWPPLKIGGSYCTSPALMHSLLRTDQMCMAAMSTMDGIDTPAVPLEPLISLTKN 240
 Db 181 IVPPLIKGSGTGPQPLDLPVPLIIGRKMVAGSAIGMKETOBMDAEEINITADIEVTPID 300
 181 IVPPLIKGSGTGPQPLDLPVPLIIGRKMVAGSAIGMKETOBMDAEEINITADIEVTPID 300
 Qy 241 GKVUTGIAVQDLPVPLIIGRKMVAGSAIGMKETOBMDAEEINITADIEVTPID 300
 241 GKVUTGIAVQDLPVPLIIGRKMVAGSAIGMKETOBMDAEEINITADIEVTPID 300
 Db 241 GKVUTGIAVQDLPVPLIIGRKMVAGSAIGMKETOBMDAEEINITADIEVTPID 300
 241 GKVUTGIAVQDLPVPLIIGRKMVAGSAIGMKETOBMDAEEINITADIEVTPID 300
 Qy 301 YINTAMERVVKDVRFRFVIDVENTL 326
 301 YINTAMERVVKDVRFRFVIDVENTL 326
 Db 301 YINTAMERVVKDVRFRFVIDVENTL 326

RESULT 3
 ADA38387 standard; protein; 355 AA.
 XX

RESULT 4	
Qy	AAB36444
Db	AAB36444 standard; protein; 333 AA.
Qy	XX
Db	XX
Qy	XX
Db	XX
DE Strawberry alcohol dehydrogenase protein SEQ ID NO:6B.	
XX	XX
KW	Strawberry; fruit flavour; biosynthetic pathway; aliphatic; thiolase;
KW	aromatic ester; alcohol acyl transferase; alcohol dehydrogenase;
KW	pyruvate decarboxylase; aminotransferase; esterase; alcohol; aldehyde;
KW	alpha-kerro acid; amino acid; fatty acid; acyl-CoA; processed food;
KW	food additive; flavour; syrup; ice-cream; frozen dessert; yoghurt;
KW	connectionary; flavouring; oral medicament; vitamin; aroma; beverage;
KW	alcohol; scent; fragrance; perfume; cosmetic; suspension aid;
KW	aluminium salt; anti-perspirant; pharmaceutical; cleaning product;
KW	insect pheromone; dye carrier; solvent; insect repellent; miticide;
KW	scabicide; plasticiser; deodorant
XX	XX
OS	Fragaria x ananassa.
XX	XX
PN	WO2000327789-A1.
XX	XX
PD	08-JUN-2000.
XX	XX
PF	02-DEC-1999;
XX	99WO-NL000737.
PR	02-DEC-1998;
PR	98EP-00204018.
PR	12-MAR-1999;
XX	99EP-003200739.
PA	(CPRO-) CPR0-CLO CENT PLANTENVERDELINGS REPROD.
XX	XX
PI	Abaroni A, Luecker J, Verhoeven HA, Van Tunen AJ, O'connell AP;
XX	XX
DR	WPI; 2000-412335/35.
DR	N-PDB; AAC6471.
XX	XX
PT	A new DNA sequence encoding a polypeptide with alcohol acyl transferase
PT	activity for producing and regulating aromatic and/or aliphatic ester
PT	formation in microorganisms, plant cells or plants.
XX	XX
PS	Claim 2B; Page 87; 163pp; English.
XX	XX
CC	The present invention describes nucleotide sequences with thiolase,
CC	alcohol acyl transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC	aminotransferase and esterase activities, which are involved in the
CC	biosynthetic pathway for aliphatic and/or aromatic ester production in
CC	fruit. The nucleotide sequences can be inserted into the genome of a
CC	fruit-producing plant to regulate aliphatic and/or aromatic ester
CC	production. Aromatic and/or aliphatic esters in microorganisms, plant
CC	cells or plants are produced by inserting thiolase, alcohol acyl
CC	transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC	aminotransferase and esterase nucleotide sequences into the genome and
CC	feeding the microorganism or plant with alcohol, aldehydes, alpha-keto
CC	acids, or amino acids and fatty acids, and acyl-CoA. The nucleotides and
CC	their proteins can be used in the processed food industry as food
CC	additives to enhance the flavour of syrups, ice-creams, frozen desserts,
CC	yoghurts and confectionery. They are used as flavouring agents for oral
CC	medications and vitamins; provide flavour and aroma in beverages,
CC	including alcohol; enhance or reduce fruit flavour, aroma, fragrance or
CC	scent; enhance the flavour of natural, synthetic or artificial flavour
CC	products; for the production of novel combinations of artificial flavour

CC substances; as antibacterial or anti-fungal agents; as fragrance or CC perfumes in cosmetics, creams, sun-protectant products, hair

CC conditioners, lengthening agents and fixatives in perfumes, suspension CC aids for aluminium salts in anti-perspirant pharmaceutical products; as

CC disinfectant additives; as personal care products; as insect CC pheromones; and as dye carriers, solvents, insect repellents, miticides, CC scabicides, plasticisers and deodorants. The present sequence represents CC the specifically claimed strawberry alcohol dehydrogenase

XX Sequence 333 AA;

Query Match 56.4%; Score 991; DB 3; Length 333;
Best Local Similarity 64.4%; Pred. No. 1.7e-92; Gaps 7;
Matches 204; Conservative 25; Mismatches 56; Indels 32; Gaps 7;

Qy 1 EITGATDVRKFLYCGVCHSDHMAANDWGTGTTSTYPTVPGHELVGVVTEVGCKYKFKF-SWR 59
Db 34 ETGEKDVTFTQMYCGICHSDHMAHMVKNEWGSTYPLVPGHELVGVTEVGSMVQKFRVDR 93

Qy 60 GCRKTDHGRPLTCNCIHLHENYCNLICGSKYDGTMTYGGSYNNMMVTDDEFIVRI 119
Db 94 VGVGCVVGSCR-SCRNCTDHLHENYCPQILLYGAKYDGTTGTYGSDIMVADEFIVRI 152

Qy 120 PDNLPLDGAAPLLCAGITTYSPWRYGLDKFGMHLGIVWPRFRSRPPLNLPGLMGSRLQ 179

Db 153 PDNLPLDGAAPLLCAGITTYSPLRYGLDKFGMHLGIVWPRFRSRPPLNLPGLMGSRLQ 195

Qy 180 SLVPPPLIKEGG---SYGTSP---ALMH---SLLRTDQDNEAAMTMDGJIDTVP 225

Db 197 HVAVKFAKAMGCVKVTWISTSKPGKEBEARKHAGDSLVSQDQMQAAIGMDGJIDTVS 256

Qy 226 AVRPPLPLISLKLTKVWVGIATQPLDVEFPLVLIIGRVAAGGKMKETQEMIDFA 285

Db 257 AQPPLPLIGLKLKGKLYMGAPEKPLEFVPLMGRKVAVGSGIGMKETQEMIDFA 316

Qy 286 AENNTADIEVITPIDYL 302

Db 317 AENNTADIEVITPIDYL 333

RESULT 5
ID AAY79661; standard; protein; 333 AA.

XX AAY79661;
XX 12-SEP-2003 (revised)
DT 29-AUG-2000 (first entry)

XX Strawberry alcohol dehydrogenase SLB39.

XX Strawberry; alcohol dehydrogenase; fruit; ripening; ester; flavour;
XX aroma; transgenic plant.

OS Fragaria x ananassa.
PN EP1006190-A1.

XX 02-DEC-1998;
XX 98EP-00204018.

XX (CPR0-) CPR0-DLO CENt PLANTENVERDELINGS REPROD.

XX Verhoeven HA, Van Tunen AJ, Abaroni A, Inecker J, O'connell AP;
XX DR WPL; 2000-37264/33-
XX N-PSDB; AAA27671.

XX PT New polynucleotides encoding enzymes from the biosynthetic pathway for
XX Populus tremoloideae.

PT aromatic and/or aliphatic ester production in fruit used to modify plant
PT flavors.

XX Claim 35; Page 79-80; 116pp; English.

CC The present sequence is that of alcohol dehydrogenase SLB39 of strawberry
CC cv. Elsanta, an enzyme involved in the metabolic pathway leading to the
CC production of volatile ester compounds in fruit, and in the ripening
CC process. The invention relates to DNA sequences (see AAU27666-78)
CC encoding enzymes (see AAU29656-68) involved in the metabolic pathway
CC leading to the formation of aliphatic and/or aromatic esters in ripening
CC fruit. The enzymes have alcohol acyl transferase, alcohol dehydrogenase,
CC pyruvate decarboxylase, thiolase or aminotransferase activity. Expression
CC vectors comprising the DNA sequences may be used to regulate ester
CC formation in fruit. Generically modified plants, plant cells and
CC microorganisms can be used to produce esters. The DNA sequences,
CC polypeptides and antibodies are also used to screen fruit: for volatile
CC ester compounds; for quality such as flavour, fragrance, aroma, scent,
CC texture or shape; to distinguish between cultivars and varieties; and to
CC monitor harvest time, post-harvest quality, shelf-life, timing of
CC pesticide application, and resistance capacity based on volatile ester
CC profiles. (Updated on 12-SEP-2003 to standardise OS field)

XX Sequence 333 AA;

Query Match 56.4%; Score 991; DB 3; Length 333;
Best Local Similarity 64.4%; Pred. No. 1.7e-92; Gaps 7;
Matches 204; Conservative 25; Mismatches 56; Indels 32; Gaps 7;

Qy 1 ETGATDVRKFLYCGVCHSDHMAANDWGTGTTSTYPTVPGHELVGVVTEVGCKYKFKF-SWR 59
Db 34 ETGEKDVTFTQMYCGICHSDHMAHMVKNEWGSTYPLVPGHELVGVTEVGSMVQKFRVDR 93

Qy 60 GCRKTDHGRPLTCNCIHLHENYCNLICGSKYDGTMTYGGSYNNMMVTDDEFIVRI 119
Db 94 VGVGCVVGSCR-SCRNCTDHLHENYCPQILLYGAKYDGTTGTYGSDIMVADEFIVRI 152

Qy 120 PDNLPLDGAAPLLCAGITTYSPWRYGLDKFGMHLGIVWPRFRSRPPLNLPGLMGSRLQ 179

Db 153 PDNLPLDGAAPLLCAGITTYSPLRYGLDKFGMHLGIVWPRFRSRPPLNLPGLMGSRLQ 195

Qy 180 SLVPPPLIKEGG---SYGTSP---ALMH---SLLRTDQDNEAAMTMDGJIDTVP 225

Db 197 HVAVKFAKAMGCVKVTWISTSKPGKEBEARKHAGDSLVSQDQMQAAIGMDGJIDTVS 256

Qy 226 AVRPPLPLISLKLTKVWVGIATQPLDVEFPLVLIIGRVAAGGKMKETQEMIDFA 285

Db 257 AQPPLPLIGLKLKGKLYMGAPEKPLEFVPLMGRKVAVGSGIGMKETQEMIDFA 316

Qy 286 AENNTADIEVITPIDYL 302

Db 317 AENNTADIEVITPIDYL 333

Query Match 56.4%; Score 991; DB 3; Length 333;
Best Local Similarity 64.4%; Pred. No. 1.7e-92; Gaps 7;
Matches 204; Conservative 25; Mismatches 56; Indels 32; Gaps 7;

Qy 1 ETGATDVRKFLYCGVCHSDHMAANDWGTGTTSTYPTVPGHELVGVVTEVGCKYKFKF-SWR 59
Db 34 ETGEKDVTFTQMYCGICHSDHMAHMVKNEWGSTYPLVPGHELVGVTEVGSMVQKFRVDR 93

Qy 60 GCRKTDHGRPLTCNCIHLHENYCNLICGSKYDGTMTYGGSYNNMMVTDDEFIVRI 119
Db 94 VGVGCVVGSCR-SCRNCTDHLHENYCPQILLYGAKYDGTTGTYGSDIMVADEFIVRI 152

Qy 120 PDNLPLDGAAPLLCAGITTYSPWRYGLDKFGMHLGIVWPRFRSRPPLNLPGLMGSRLQ 179

Db 153 PDNLPLDGAAPLLCAGITTYSPLRYGLDKFGMHLGIVWPRFRSRPPLNLPGLMGSRLQ 195

Qy 180 SLVPPPLIKEGG---SYGTSP---ALMH---SLLRTDQDNEAAMTMDGJIDTVP 225

Db 197 HVAVKFAKAMGCVKVTWISTSKPGKEBEARKHAGDSLVSQDQMQAAIGMDGJIDTVS 256

Qy 226 AVRPPLPLISLKLTKVWVGIATQPLDVEFPLVLIIGRVAAGGKMKETQEMIDFA 285

Db 257 AQPPLPLIGLKLKGKLYMGAPEKPLEFVPLMGRKVAVGSGIGMKETQEMIDFA 316

Qy 286 AENNTADIEVITPIDYL 302

Db 317 AENNTADIEVITPIDYL 333

RESULT 6
ID AAU0013 standard; protein; 362 AA.

XX AAU0013;

XX 15-JUL-2002 (first entry)

XX Sinapyl alcohol dehydrogenase (SAD).

XX Plant; aspen; phenylpropanoid pathway; agronomic; lignin; paper;

XX 4-coumarate-CoA ligase; 4CBL; coniferyl alcohol 5-hydroxylase; CA16EH;

XX S-adenosyl-L-methionine-dependent; SAM; AldoMT; transgenic; grass;

XX 5-hydroxyconiferaldehyde O-methyltransferase; cellulose; pulp;

XX syringyl; guaiacyl; agriculture; CAD; sinapyl alcohol dehydrogenase; SAD;

XX OS Populus tremoloideae.

PR 25-FEB-1999; 990US-0121825P; 990US-0143542P.
 PR 09-MAR-1999; 990US-0123180P; 990US-0144005P.
 PR 23-MAR-1999; 990US-0123548P; 990US-0144052P.
 PR 29-MAR-1999; 990US-0125788P; 990US-0144085P.
 PR 01-APR-1999; 990US-01262264P; 990US-0144086P.
 PR 08-APR-1999; 990US-0127462P; 990US-0144325P.
 PR 16-APR-1999; 990US-0128234P; 990US-0144332P.
 PR 21-APR-1999; 990US-0130077P; 990US-0144333P.
 PR 23-APR-1999; 990US-0130449P; 990US-0144335P.
 PR 28-APR-1999; 990US-01310510P; 990US-0144335P.
 PR 30-APR-1999; 990US-01312407P; 990US-0144335P.
 PR 04-MAY-1999; 990US-0131324B4P; 990US-0144335P.
 PR 05-MAY-1999; 990US-0131235P; 990US-0144335P.
 PR 06-MAY-1999; 990US-01313449P; 990US-0144335P.
 PR 07-MAY-1999; 990US-01312487P; 990US-0144335P.
 PR 14-MAY-1999; 990US-01314256P; 990US-0144335P.
 PR 14-MAY-1999; 990US-01314218P; 990US-0144335P.
 PR 14-MAY-1999; 990US-01314219P; 990US-0144335P.
 PR 14-MAY-1999; 990US-01314221P; 990US-0144335P.
 PR 18-MAY-1999; 990US-01314768P; 990US-0144335P.
 PR 20-MAY-1999; 990US-01315124P; 990US-0144335P.
 PR 21-MAY-1999; 990US-01315353P; 990US-0144335P.
 PR 25-MAY-1999; 990US-01316021P; 990US-0144335P.
 PR 27-MAY-1999; 990US-01316392P; 990US-0144335P.
 PR 01-JUN-1999; 990US-01316782P; 990US-0144335P.
 PR 03-JUN-1999; 990US-01317222P; 990US-0144335P.
 PR 14-JUN-1999; 990US-01317529E; 990US-0144335P.
 PR 16-JUN-1999; 990US-01317502E; 990US-0144335P.
 PR 07-JUN-1999; 990US-0137724P; 990US-0144335P.
 PR 08-JUN-1999; 990US-0138094E; 990US-0144335P.
 PR 10-JUN-1999; 990US-0138540P; 990US-0144335P.
 PR 18-JUN-1999; 990US-0138847P; 990US-0144335P.
 PR 18-JUN-1999; 990US-0139119P; 990US-0144335P.
 PR 18-JUN-1999; 990US-0139452P; 990US-0144335P.
 PR 16-JUN-1999; 990US-0139453P; 990US-0144335P.
 PR 17-JUN-1999; 990US-0139492P; 990US-0144335P.
 PR 18-JUN-1999; 990US-0139454D; 990US-0144335P.
 PR 18-JUN-1999; 990US-0139455P; 990US-0144335P.
 PR 18-JUN-1999; 990US-0139456P; 990US-0144335P.
 PR 18-JUN-1999; 990US-0139462P; 990US-0144335P.
 PR 18-JUN-1999; 990US-0139463P; 990US-0144335P.
 PR 18-JUN-1999; 990US-0139458P; 990US-0144335P.
 PR 18-JUN-1999; 990US-0139750P; 990US-0144335P.
 PR 21-JUN-1999; 990US-0139763P; 990US-0144335P.
 PR 22-JUN-1999; 990US-0139817P; 990US-0144335P.
 PR 23-JUN-1999; 990US-0139460P; 990US-0144335P.
 PR 23-JUN-1999; 990US-0140353P; 990US-0144335P.
 PR 23-JUN-1999; 990US-0141842P; 990US-0144335P.
 PR 01-JUL-1999; 990US-0142154P; 990US-0144335P.
 PR 02-JUL-1999; 990US-0142055P; 990US-0144335P.
 PR 06-JUL-1999; 990US-0142390P; 990US-0155659P.
 PR 08-JUL-1999; 990US-0142803P; 990US-0155458P.
 PR 09-JUL-1999; 990US-0142920P; 990US-0156596P.
 PR 12-JUL-1999; 990US-0142977P; 990US-0156596P.

PR	18-JUN-1999;	990US-0139462P.	PR	25-AUG-1999;	990US-0150566P.
PR	18-JUN-1999;	990US-0139463P.	PR	26-AUG-1999;	990US-0150584P.
PR	18-JUN-1999;	990US-0139750P.	PR	27-AUG-1999;	990US-0151065P.
PR	18-JUN-1999;	990US-0139763P.	PR	27-AUG-1999;	990US-0151066P.
PR	21-JUN-1999;	990US-0139817P.	PR	27-AUG-1999;	990US-0151080P.
PR	22-JUN-1999;	990US-0139899P.	PR	30-AUG-1999;	990US-0151303P.
PR	23-JUN-1999;	990US-0140353P.	PR	31-AUG-1999;	990US-0151438P.
PR	23-JUN-1999;	990US-0140544P.	PR	01-SEP-1999;	990US-0151930P.
PR	24-JUN-1999;	990US-0140695P.	PR	01-SEP-1999;	990US-0152163P.
PR	28-JUN-1999;	990US-0140823P.	PR	07-SEP-1999;	990US-0153070P.
PR	29-JUN-1999;	990US-0140911P.	PR	10-SEP-1999;	990US-0153079P.
PR	30-JUN-1999;	990US-0141287P.	PR	13-SEP-1999;	990US-0153158P.
PR	01-JUL-1999;	990US-0141842P.	PR	15-SEP-1999;	990US-0154018P.
PR	01-JUL-1999;	990US-0142154P.	PR	16-SEP-1999;	990US-0154039P.
PR	02-JUL-1999;	990US-0142055P.	PR	20-SEP-1999;	990US-0154717P.
PR	06-JUL-1999;	990US-0142390P.	PR	04-OCT-1999;	990US-0157117P.
PR	08-JUL-1999;	990US-0142803P.	PR	05-OCT-1999;	990US-0157753P.
PR	09-JUL-1999;	990US-0142220P.	PR	06-OCT-1999;	990US-0157865P.
PR	12-JUL-1999;	990US-0142977P.	PR	07-OCT-1999;	990US-0158029P.
PR	13-JUL-1999;	990US-0143542P.	PR	08-OCT-1999;	990US-0158322P.
PR	14-JUL-1999;	990US-0143624P.	PR	09-OCT-1999;	990US-0158369P.
PR	15-JUL-1999;	990US-0144005P.	PR	10-OCT-1999;	990US-0158390P.
PR	16-JUL-1999;	990US-0144086P.	PR	11-OCT-1999;	990US-0158391P.
PR	17-JUL-1999;	990US-0144334P.	PR	12-OCT-1999;	990US-0158392P.
PR	19-JUL-1999;	990US-0144334P.	PR	13-OCT-1999;	990US-0158393P.
PR	19-JUL-1999;	990US-0144335P.	PR	14-OCT-1999;	990US-0159310P.
PR	20-JUL-1999;	990US-0144335P2.	PR	14-OCT-1999;	990US-0159311P.
PR	20-JUL-1999;	990US-01446332P.	PR	14-OCT-1999;	990US-0159637P.
PR	20-JUL-1999;	990US-0144884P.	PR	14-OCT-1999;	990US-0159688P.
PR	21-JUL-1999;	990US-0144814P.	PR	18-OCT-1999;	990US-015984P.
PR	21-JUL-1999;	990US-0145086P.	PR	21-OCT-1999;	990US-0159741P.
PR	21-JUL-1999;	990US-0145088P.	PR	14-OCT-1999;	990US-0160176P.
PR	22-JUL-1999;	990US-0145085P.	PR	21-OCT-1999;	990US-0160776P.
PR	22-JUL-1999;	990US-0145087P.	PR	21-OCT-1999;	990US-0160770P.
PR	22-JUL-1999;	990US-0145098P.	PR	21-OCT-1999;	990US-0160814P.
PR	22-JUL-1999;	990US-0145192P.	PR	25-OCT-1999;	990US-0161405P.
PR	23-JUL-1999;	990US-0145145P.	PR	22-OCT-1999;	990US-0161698P.
PR	23-JUL-1999;	990US-0145218P.	PR	26-OCT-1999;	990US-0161359P.
PR	23-JUL-1999;	990US-0145224P.	PR	26-OCT-1999;	990US-0161360P.
PR	24-JUL-1999;	990US-01452276P.	PR	28-OCT-1999;	990US-0161920P.
PR	27-JUL-1999;	990US-0145913P.	PR	28-OCT-1999;	990US-0161920P.
PR	27-JUL-1999;	990US-0145918P.	PR	28-OCT-1999;	990US-0161933P.
PR	28-JUL-1999;	990US-0145919P.	PR	29-OCT-1999;	990US-0162142P.
PR	02-AUG-1999;	990US-0145951P.			
PR	02-AUG-1999;	990US-0146362P.			
PR	02-AUG-1999;	990US-0146388P.			
PR	03-AUG-1999;	990US-0147204P.			
PR	04-AUG-1999;	990US-0147302P.			
PR	05-AUG-1999;	990US-0147192P.			
PR	06-AUG-1999;	990US-0147260P.			
PR	06-AUG-1999;	990US-0147303P.			
PR	06-AUG-1999;	990US-0147416P.			
PR	09-AUG-1999;	990US-0147493P.			
PR	10-AUG-1999;	990US-0148171P.			
PR	11-AUG-1999;	990US-0148319P.			
PR	12-AUG-1999;	990US-0148341P.			
PR	13-AUG-1999;	990US-0148565P.			
PR	13-AUG-1999;	990US-0148684P.			
PR	16-AUG-1999;	990US-0149368P.			
PR	17-AUG-1999;	990US-0149175P.			
PR	18-AUG-1999;	990US-0149426P.			
PR	20-AUG-1999;	990US-0149722P.			
PR	20-AUG-1999;	990US-0149723P.			
PR	23-AUG-1999;	990US-0149902P.			
PR	23-AUG-1999;	990US-0149930P.			

Query	Match	50.1%; Score 81; DB 3; Length 360;
Query	Best Local Similarity 49.1%; Pred. No. 3.6-81; Mismatches 77; Inverts 50; Gaps 5;	Best Local Similarity 49.1%; Pred. No. 3.6-81; Mismatches 71; Inverts 50; Gaps 5;
Query	1 ETGATDVRFXVLYCGYCHSDIHMKNWDGTSYPIVPGHFLVSYTEVGCKVKFKPSWQ 60	1 ETGATDVRFXVLYCGYCHSDIHMKNWDGTSYPIVPGHFLVSYTEVGCKVKFKPSWQ 60
Db	34 DNGENDTVKILFCVHTDLHTKNDWGYSSYVPGHBEIVGATKVNKVTKFKEGDR 93	34 DNGENDTVKILFCVHTDLHTKNDWGYSSYVPGHBEIVGATKVNKVTKFKEGDR 93
Query	61 1F701P0N1PLDGAFLGAGTTPSMPRYGLDJKPGRHLSWMPRFRSPPLNLPLGL 113	61 1F701P0N1PLDGAFLGAGTTPSMPRYGLDJKPGRHLSWMPRFRSPPLNLPLGL 113
Db	94 VGVGVISGSC-----QSCESCDQDLNENYCPMSFTYNAIGSGCTKNGYSENIVDQ 146	94 VGVGVISGSC-----QSCESCDQDLNENYCPMSFTYNAIGSGCTKNGYSENIVDQ 146
Query	114 HFTORIPDNLPLDGAFLGAGTTPSMPRYGLDJKPGRHLSWMPRFRSPPLNLPLGL 173	114 HFTORIPDNLPLDGAFLGAGTTPSMPRYGLDJKPGRHLSWMPRFRSPPLNLPLGL 173
Db	147 RPVDRPENIPSDGASPLLAGTIVYSPKIGTBAEGLW-----AGL 192	147 RPVDRPENIPSDGASPLLAGTIVYSPKIGTBAEGLW-----AGL 192
Query	174 WGSRLQSLLYPLIKEGGSYGTSPALMHS-----LRTDODMERAAMST 216	174 WGSRLQSLLYPLIKEGGSYGTSPALMHS-----LRTDODMERAAMST 216
Db	193 GG-----LGHVAVXKGAFGLKVTVISSSSKAERINLGHADSLFLVTTDQRMKAIGT 247	193 GG-----LGHVAVXKGAFGLKVTVISSSSKAERINLGHADSLFLVTTDQRMKAIGT 247

Qy	217	MDGIIIDTVPAVREPLISSLLKTKNGKVVTVGIAVOPDLDLWSPFLIGRKNVAGSATGGMK	276	WGSRLQSLVPLPLIKRGGSGSYGSPALMHS-----LURTDQDOMEAAMST 216
Db	248	MDYIDITSAATHPLLGLKVNGLLALSLPERLELPMFLVGRKNGGSDVGGMK	307	193 GS---LGHAVKICKPLGKIVTIVSSSTKAEEAANHGLADSPLVTTDQKMKAGIT 247
Qy	277	ETQEMIDPAEHNNTADIEVTPIDVINTAMERVVKDVRERFVIDVENTI	326	217 MDGIIIDTVPAVREPLISSLLKTKNGKVVTVGIAVOPDLDLWSPFLIGRKNVAGSATGGMK 276
Db	308	ETQEMIDFCARHNTADIEVTPIDVINTAMERLAKSDVRERFVIDVANSL	357	248 MDYIDITSAATHPLLGLKVNGLLALSLPERLELPMFLVGRKNGGSDVGGMK 307
RESUL	13			
ABB93303		standard; protein; 360 AA.		
XX				
AC	ABB93303;			
XX				
DT	31-MAY-2002	(first entry)		
XX				
DE	Herbicidally active polypeptide	SEQ ID NO 2514.		
XX				
KW	Herbicide; Plant; agriculture; herbicide.			
XX				
OS	Arabidopsis thaliana.			
XX				
PN	WO200210210-A2.			
XX				
PD	07-FEB-2002.			
XX				
PF	28-AUG-2001; 2001WO-EP0009892.			
XX				
PR	28-AUG-2001; 2001WO-EP0009892.			
XX				
PA	{FARB } BAYER AG.			
XX				
PI	Tietjen K, Weidler M;			
XX				
WPI;	2002-269010/31.			
DR				
XX				
PT	Identifying plant target proteins for herbicidally active compounds.			
PT	comprising aligning and comparing nucleic acid or amino acid sequences			
PT	from plant with nucleic acid or amino acid sequences from non-plant			
PT	organisms.			
XX				
PS	SEQ ID NO 2514; 261pp + Sequence Listing: English.			
XX				
CC	The invention relates to identifying target proteins (ABB90790-ABB941616)			
CC	for herbicidally active compounds, comprising aligning and comparing			
CC	nucleic acid or amino acid sequences from plant with nucleic acid or			
CC	amino acid sequences from non-plant organisms using suitable search			
CC	parameters, where plant sequences having an E-value greater than			
CC	3 of 3 than the E-value of most similar non-plant sequences are selected.			
CC	The polypeptides or nucleic acids encoding them are useful for			
CC	identifying modulators. The identified modulators are useful as			
XX				
SQ	Sequence 360 AA;			
Qy	50.1%;	Score 881; DB 5; Length 360;		
Best Local Similarity	49.1%;	Pred. No. 3.6e-81;		
Matches 172;	Conservative 51;	Mismatches 77; Indels 50; Gaps 5;		
Qy	1	ETGATDVRFRKTYCGYCHSDIHMARDWGTTSTYPTPGHFLVGVTEVGKVKRKFWSRQ	60	
Db	34	DNGENDVTVKLLFCGTYHTDHTTNDGYYYPVPGHFLVGVTEVGKVKRKFWSRQ	93	
Qy	61	-----GRCWLRGLRPTCNCIHLNCTYGSXYDGMITYEGGSNNNNVTDIE 113		
Db	94	VGVGVISGSC-----QSCESCDQDLENCPQMSFTYNAIGSDGTGKNTGEGYSNTVVDQ	146	
Qy	114	HFIYVLPDNFLPLDGAAPLICAGITTYSPMYYGJDKPGMELIGVENVRRFSRPPNLPGI	173	
Db	147	RFVLRFPENLPSDGAAPLICAGITTYSPMYYGJDKPGMELIGVENVRRFSRPPNLPGI-----AGI	192	

PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.	PR	09-AUG-1999;	99US-0147433P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-014793P.	PR	10-AUG-1999;	99US-014811P.
PR	10-JUN-1999;	99US-0138540P.	PR	10-AUG-1999;	99US-014811P.	PR	11-AUG-1999;	99US-0148319P.
PR	11-JUN-1999;	99US-0138841P.	PR	12-AUG-1999;	99US-014831P.	PR	13-AUG-1999;	99US-014855P.
PR	14-JUN-1999;	99US-0139119P.	PR	13-AUG-1999;	99US-014864P.	PR	13-AUG-1999;	99US-014864P.
PR	16-JUN-1999;	99US-0139452P.	PR	18-JUN-1999;	99US-013945P.	PR	16-AUG-1999;	99US-014938P.
PR	16-JUN-1999;	99US-0139453P.	PR	18-JUN-1999;	99US-013945P.	PR	17-AUG-1999;	99US-014915P.
PR	17-JUN-1999;	99US-0139452P.	PR	18-JUN-1999;	99US-013945P.	PR	18-AUG-1999;	99US-014946P.
PR	18-JUN-1999;	99US-0139453P.	PR	18-JUN-1999;	99US-013945P.	PR	20-AUG-1999;	99US-014972P.
PR	18-JUN-1999;	99US-0139458P.	PR	18-JUN-1999;	99US-013945P.	PR	20-AUG-1999;	99US-014973P.
PR	18-JUN-1999;	99US-0139459P.	PR	18-JUN-1999;	99US-013945P.	PR	20-AUG-1999;	99US-014995P.
PR	18-JUN-1999;	99US-0139460P.	PR	21-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-014992P.
PR	18-JUN-1999;	99US-0139461P.	PR	18-JUN-1999;	99US-0139462P.	PR	23-AUG-1999;	99US-014993P.
PR	18-JUN-1999;	99US-0139462P.	PR	18-JUN-1999;	99US-0139463P.	PR	25-AUG-1999;	99US-015056P.
PR	18-JUN-1999;	99US-0139463P.	PR	18-JUN-1999;	99US-0139464P.	PR	26-AUG-1999;	99US-015084P.
PR	18-JUN-1999;	99US-0139464P.	PR	18-JUN-1999;	99US-0139465P.	PR	27-AUG-1999;	99US-015105P.
PR	18-JUN-1999;	99US-0139465P.	PR	18-JUN-1999;	99US-0139466P.	PR	27-AUG-1999;	99US-015106P.
PR	18-JUN-1999;	99US-0139466P.	PR	19-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-015107P.
PR	22-JUN-1999;	99US-0139817P.	PR	22-JUN-1999;	99US-0139818P.	PR	27-AUG-1999;	99US-015108P.
PR	23-JUN-1999;	99US-0139818P.	PR	23-JUN-1999;	99US-0139819P.	PR	30-AUG-1999;	99US-015130P.
PR	23-JUN-1999;	99US-0139819P.	PR	23-JUN-1999;	99US-0139820P.	PR	31-AUG-1999;	99US-015143P.
PR	24-JUN-1999;	99US-0139750P.	PR	24-JUN-1999;	99US-0139751P.	PR	25-AUG-1999;	99US-015143P.
PR	28-JUN-1999;	99US-0139753P.	PR	28-JUN-1999;	99US-0139754P.	PR	26-AUG-1999;	99US-015143P.
PR	29-JUN-1999;	99US-0140991P.	PR	29-JUN-1999;	99US-0140992P.	PR	27-SEP-1999;	99US-015223P.
PR	30-JUN-1999;	99US-0140992P.	PR	30-JUN-1999;	99US-0141287P.	PR	07-SEP-1999;	99US-015233P.
PR	01-JUL-1999;	99US-0141287P.	PR	01-JUL-1999;	99US-0141842P.	PR	10-SEP-1999;	99US-015307P.
PR	01-JUL-1999;	99US-0141842P.	PR	01-JUL-1999;	99US-0142154P.	PR	10-SEP-1999;	99US-015307P.
PR	02-JUL-1999;	99US-0142054P.	PR	02-JUL-1999;	99US-0140693P.	PR	13-SEP-1999;	99US-015373P.
PR	06-JUL-1999;	99US-0142360P.	PR	06-JUL-1999;	99US-0140821P.	PR	13-SEP-1999;	99US-015373P.
PR	08-JUL-1999;	99US-0142803P.	PR	08-JUL-1999;	99US-0142804P.	PR	13-SEP-1999;	99US-015373P.
PR	12-JUL-1999;	99US-0142970P.	PR	12-JUL-1999;	99US-0142971P.	PR	15-SEP-1999;	99US-015408P.
PR	13-JUL-1999;	99US-0143542P.	PR	13-JUL-1999;	99US-0143543P.	PR	16-SEP-1999;	99US-015409P.
PR	14-JUL-1999;	99US-0143624P.	PR	14-JUL-1999;	99US-0144005P.	PR	01-SEP-1999;	99US-0154779P.
PR	15-JUL-1999;	99US-0143624P.	PR	15-JUL-1999;	99US-0144331P.	PR	02-SEP-1999;	99US-0155129P.
PR	16-JUL-1999;	99US-0144005P.	PR	16-JUL-1999;	99US-0144006P.	PR	10-SEP-1999;	99US-0155486P.
PR	16-JUL-1999;	99US-0144006P.	PR	16-JUL-1999;	99US-0144293P.	PR	13-SEP-1999;	99US-0155648P.
PR	19-JUL-1999;	99US-0144325P.	PR	19-JUL-1999;	99US-0144326P.	PR	29-SEP-1999;	99US-015656P.
PR	20-JUL-1999;	99US-0144331P.	PR	20-JUL-1999;	99US-0144332P.	PR	04-OCT-1999;	99US-015717P.
PR	20-JUL-1999;	99US-0144332P.	PR	20-JUL-1999;	99US-0144333P.	PR	05-OCT-1999;	99US-0157753P.
PR	21-JUL-1999;	99US-0144333P.	PR	21-JUL-1999;	99US-0144334P.	PR	06-OCT-1999;	99US-015785P.
PR	19-JUL-1999;	99US-0144334P.	PR	19-JUL-1999;	99US-0144335P.	PR	24-SEP-1999;	99US-015809P.
PR	19-JUL-1999;	99US-0144335P.	PR	19-JUL-1999;	99US-0144336P.	PR	08-OCT-1999;	99US-015822P.
PR	20-JUL-1999;	99US-0144336P.	PR	20-JUL-1999;	99US-0144370P.	PR	12-OCT-1999;	99US-015839P.
PR	20-JUL-1999;	99US-0144370P.	PR	20-JUL-1999;	99US-0144632P.	PR	13-OCT-1999;	99US-015923P.
PR	21-JUL-1999;	99US-0144632P.	PR	21-JUL-1999;	99US-0144634P.	PR	13-OCT-1999;	99US-015924P.
PR	21-JUL-1999;	99US-0144634P.	PR	21-JUL-1999;	99US-0144635P.	PR	14-OCT-1999;	99US-015933P.
PR	19-JUL-1999;	99US-0144635P.	PR	19-JUL-1999;	99US-0144636P.	PR	14-OCT-1999;	99US-015933P.
PR	22-JUL-1999;	99US-0144636P.	PR	22-JUL-1999;	99US-0144637P.	PR	21-OCT-1999;	99US-015967P.
PR	22-JUL-1999;	99US-0144637P.	PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-016084P.
PR	22-JUL-1999;	99US-0145089P.	PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-016084P.
PR	23-JUL-1999;	99US-0145145P.	PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-016091P.
PR	22-JUL-1999;	99US-0145218P.	PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-016091P.
PR	26-JUL-1999;	99US-0145224P.	PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-016144P.
PR	27-JUL-1999;	99US-0145276P.	PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-016145P.
PR	27-JUL-1999;	99US-0145913P.	PR	27-JUL-1999;	99US-0145918P.	PR	26-OCT-1999;	99US-016159P.
PR	27-JUL-1999;	99US-0145918P.	PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-016160P.
PR	02-AUG-1999;	99US-0146366P.	PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-016161P.
PR	02-AUG-1999;	99US-0146386P.	PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-016192P.
PR	02-AUG-1999;	99US-0146388P.	PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-016192P.
PR	03-AUG-1999;	99US-0147038P.	PR	03-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147204P.	PR	04-AUG-1999;	99US-0147302P.	PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147302P.	PR	05-AUG-1999;	99US-0147303P.	PR	05-AUG-1999;	99US-0147303P.

Query Match 49.8%; Score 876; DB 3; Length 360;
 Best Local Similarity 48.9%; Pred. No. 1.2e-80;
 Matches 171; Conservative 52; Mismatches 77; Indels 50; Gaps 5;

Qy	1	STGATDVRKFVLYCCSYCHSD1HMAKNDWGTSTTYPIVPGHELNGVATVPGCKVKKFKSWRQ	60
	:	: : : : : : : : : : : : : : : : :	:
Db	34	DNGENDVTVKILFCGVCHTDLHNTNDWGXYYPPVPGHETVGIATKVRGNVTKFKEGDR	93
Qy	61	-----GRCLHLGRURPTCENCIHLENYCPNLQITYGSKYDGMITYGGYSNNWVTD	113
Db	94	VGVGVTGSGC-----QSCESCDQDILENYCPOMSFTYNAIGSDDGTRNGYSENIIVDQ	146
Qy	114	HFIYR1PDNLPLDGAAPLLCAGITIYSPWRYGLDRKGMLHGVIEWPFRSRSPPLNPGL	173
Db	147	RFVLRPPENL2SDSAPLLCAGITIYSPMKYXGMITAGKHLGV-----	192
Qy	174	WGSR1QSLVPPPLKRGSGYTSPLMHS-----	216
Db	193	GG-----LGEWAKVKGKAFGLKVTISSLSSSTKAEAINHLAGDSFLVTDPKRKAIGT	247
Qy	217	MDG1IDTVDPYRPLPPLISLKTNGKNTVYGLAQVQLDLPVPLILGRKOMYGA1GSMK	276
Db	248	MDY1DTTISAVHALYPLPLGLIKVNGKLIALGLPEKPLDPMFPLVLRKOMYGSDDVCSMK	307
Qy	277	ETQEMIDFAEHNNTADIEVTPIDVLTNTAERVVVKDVRFREV1DVENTL	326
Db	308	ETQEMIDFCAKHNNTADIELJKMDDENTRALASSDVRYRFVINTVANSL	357
RESULT 15			
		AAG29471	standard; prototin; 306 AA.
	ID	AAG29471	
	XX	XX	
	AC	XX	
	XX	XX	
	DT	17-OCT-2000	(first entry)
	XX	XX	
	DB	Arabidopsis thaliana	Protein fragment SEQ ID NO: 35072.
	XX	XX	
	KW	Protein identification; signal transduction pathway; metabolic pathway;	
	KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
	KW	termination sequence.	
	XX	XX	
	OS	Arabidopsis thaliana.	
	PN	EP1033405-A2.	
	XX		
	PD	06-SEP-2000.	
	XX		
	PP	25-FEB-2000; 20000EP-00101439.	
	XX		
	PR	25-FEB-1999;	
	PR	05-MAR-1999;	99US-0121825P.
	PR	09-MAR-1999;	99US-0123180P.
	PR	23-MAR-1999;	99US-01231548P.
	PR	25-MAR-1999;	99US-0125788P.
	PR	29-MAR-1999;	99US-0126264P.
	PR	01-APR-1999;	99US-0126785P.
	PR	06-APR-1999;	99US-0127462P.
	PR	08-APR-1999;	99US-0128234P.
	PR	16-APR-1999;	99US-0128714P.
	PR	19-APR-1999;	99US-0129845P.
	PR	21-APR-1999;	99US-0130077P.
	PR	04-MAY-1999;	99US-0130449P.
	PR	23-APR-1999;	99US-0130510P.
	PR	06-MAY-1999;	99US-0130891P.
	PR	28-APR-1999;	99US-0131449P.
	PR	30-APR-1999;	99US-013204BP.
	PR	30-APR-1999;	99US-0132407P.
	PR	04-MAY-1999;	99US-0132484P.
	PR	05-MAY-1999;	99US-0132485P.
	PR	06-MAY-1999;	99US-0132486P.
	PR	07-MAY-1999;	99US-0132487P.
	PR	11-MAY-1999;	99US-0132863P.
	PR	14-MAY-1999;	99US-0134256P.
	PR	14-MAY-1999;	99US-0134218P.
	PR	14-MAY-1999;	99US-0134219P.

PR	14-MAY-1999;	990US-0134221P.
PR	14-MAY-1999;	990US-01346168P.
PR	18-MAY-1999;	990US-0134941P.
PR	19-MAY-1999;	990US-0135124P.
PR	21-MAY-1999;	990US-0135535P.
PR	24-MAY-1999;	990US-0135679P.
PR	25-MAY-1999;	990US-0136202P.
PR	27-MAY-1999;	990US-0136592P.
PR	08-MAY-1999;	990US-0136782P.
PR	01-JUN-1999;	990US-0137222P.
PR	03-JUN-1999;	990US-0137228P.
PR	04-JUN-1999;	990US-0137502P.
PR	07-JUN-1999;	990US-0137724P.
PR	08-JUN-1999;	990US-0138094P.
PR	10-JUN-1999;	990US-0138540P.
PR	10-JUN-1999;	990US-0138847P.
PR	14-JUN-1999;	990US-0139119P.
PR	16-JUN-1999;	990US-0139452P.
PR	16-JUN-1999;	990US-0139453P.
PR	17-JUN-1999;	990US-0139492P.
PR	18-JUN-1999;	990US-0139454P.
PR	18-JUN-1999;	990US-0139455P.
PR	18-JUN-1999;	990US-0139456P.
PR	18-JUN-1999;	990US-0139457P.
PR	18-JUN-1999;	990US-0139458P.
PR	18-JUN-1999;	990US-0139459P.
PR	18-JUN-1999;	990US-0139460P.
PR	18-JUN-1999;	990US-0139461P.
PR	18-JUN-1999;	990US-0139462P.
PR	18-JUN-1999;	990US-0139463P.
PR	18-JUN-1999;	990US-0139750P.
PR	18-JUN-1999;	990US-0139763P.
PR	21-JUN-1999;	990US-0139817P.
PR	22-JUN-1999;	990US-0139899P.
PR	23-JUN-1999;	990US-0140353P.
PR	23-JUN-1999;	990US-01419463P.
PR	24-JUN-1999;	990US-0142055P.
PR	28-JUN-1999;	990US-0142390P.
PR	29-JUN-1999;	990US-0142803P.
PR	30-JUN-1999;	990US-0142920P.
PR	01-JUL-1999;	990US-0142927P.
PR	01-JUL-1999;	990US-0141842P.
PR	02-JUL-1999;	990US-0141544P.
PR	06-JUL-1999;	990US-0142055P.
PR	08-JUL-1999;	990US-0142803P.
PR	09-JUL-1999;	990US-0142920P.
PR	12-JUL-1999;	990US-0142927P.
PR	13-JUL-1999;	990US-0143542P.
PR	14-JUL-1999;	990US-0143624P.
PR	15-JUL-1999;	990US-0144005P.
PR	16-JUL-1999;	990US-0144085P.
PR	16-JUL-1999;	990US-0144086P.
PR	16-JUL-1999;	990US-0144225P.
PR	19-JUL-1999;	990US-0144331P.
PR	19-JUL-1999;	990US-0144332P.
PR	19-JUL-1999;	990US-0144334P.
PR	21-JUL-1999;	990US-0145086P.
PR	21-JUL-1999;	990US-0145088P.
PR	22-JUL-1999;	990US-0145085P.
PR	22-JUL-1999;	990US-0145087P.
PR	22-JUL-1999;	990US-0145089P.
PR	22-JUL-1999;	990US-0145192P.
PR	23-JUL-1999;	990US-0145145P.
PR	23-JUL-1999;	990US-0145224P.
PR	26-JUL-1999;	990US-0145276P.

PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161399P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161301P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161301P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161900P.
PR	03-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161902P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147302P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	07-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147433P.			
PR	09-AUG-1999;	99US-0147433P.			
PR	10-AUG-1999;	99US-0148170P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148555P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	26-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150844P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151203P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151530P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0155375P.			
PR	15-SEP-1999;	99US-0155401P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0154779P.			
PR	23-SEP-1999;	99US-0155139P.			
PR	24-SEP-1999;	99US-0155486P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	08-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0155559P.			
PR	08-OCT-1999;	99US-015829P.			
PR	12-OCT-1999;	99US-0158322P.			
PR	13-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159393P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	14-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	18-OCT-1999;	99US-0159594P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			
PR	21-OCT-1999;	99US-0160814P.			
PR	22-OCT-1999;	99US-0160980P.			
PR	22-OCT-1999;	99US-0160981P.			
PR	25-OCT-1999;	99US-0161404P.			

Query Match 49.5%; Score 871; DB 3; Length 306;
 Best Local Similarity 54.7%; Pred. No. 3e-80; Mismatches 65; Indels 38; Gaps 5;

Matches 176; Conservative 43; Job time : 44.3213 secs

PR	23	MAKNDEUTSTPIVPLGEBLWGYTYE/CCKVKKFKSW-ROGRCMTHGFLRPTCENC/CHLE 81
	1	MARNEGDITVPLGEBLWGYTYE/CCKVKKFKSW-SCDSNCNGDE 59
QY	82	NYCPNL1QTYGSKYDOTMTYGGSNMVTDEHFTVIRPDNEVPLGAPLCAIGITYSP 141
Db	60	NYCPNL1QTYGSKYDOTMTYGGSNMVTDEHFTVIRPDNEVPLGAPLCAIGITYSP 119
QY	245	TUGIAVQPLDLPVFPLIGREKVGASAIIGMKEKTCMDEAHEHTTADIEVTPIDYLN 304
Db	221	WRYYGLDKPGMELGVENPFRFRSPPLNPGWRSRQSLVPLIKEGGSYTSPALMHS 201
QY	305	AMERVVKDVKRFRFVIDVENTL 326
Db	281	AMERLAKADVKYRFVIDVANTM 302
QY	202	LLRTDQDNEAAMSMDGIDTVAFLGADFLVSDEKQMDAMTMDGIDTWSATHPLLPGILKNGKLV 244
Db	161	SERRDDEAVTRLJADFLVSLVSEKQMDAMTMDGIDTWSATHPLLPGILKNGKLV 220
QY	245	TUGIAVQPLDLPVFPLIGREKVGASAIIGMKEKTCMDEAHEHTTADIEVTPIDYLN 304
Db	221	WYGAAPPLDPVFLPFLIGREKVGASNIVG3IKETQEMVNDLACKHNTTADIELISADYVNT 280

Search completed: July 6, 2004, 13:36:45

Result No.	Score	Query Match	Length	DB ID	Description
1	1007	57.3	355	4 US-09-501-115-12	Sequence 12, Appl
2	869	49.4	360	4 US-09-501-115-4	Sequence 4, Appl
3	832	47.3	371	4 US-09-501-115-14	Sequence 14, Appl
4	800	45.5	361	4 US-09-501-115-8	Sequence 8, Appl
5	790	44.9	366	4 US-09-501-115-10	Sequence 10, Appl
6	732.5	41.7	364	4 US-09-501-115-2	Sequence 2, Appl
7	717	40.8	358	4 US-09-501-115-6	Sequence 6, Appl
8	665	37.8	360	4 US-09-501-115-12	Sequence 32, Appl
9	648	36.9	358	4 US-09-501-115-30	Sequence 30, Appl
10	630.5	35.9	353	4 US-09-328-352-7473	Sequence 7473, Appl
11	577.5	33.0	410	4 US-09-252-9191A-25647	Sequence 25647, Appl
12	578	32.9	357	4 US-09-501-115-20	Sequence 20, Appl
13	567	32.3	358	4 US-09-489-0394-10508	Sequence 10508, Appl
14	559	31.8	329	4 US-09-501-115-28	Sequence 28, Appl
15	379	21.6	360	4 US-09-501-115-34	Sequence 34, Appl
16	366.5	20.8	374	4 US-09-489-0394-8925	Sequence 8925, Appl
17	356	20.3	358	4 US-09-328-352-4988	Sequence 4988, Appl
18	316	18.0	344	4 US-09-134-001C-5343	Sequence 5343, Appl
19	310	17.6	349	4 US-07-857-24B-95	Sequence 95, Appl
20	303	17.2	186	4 US-09-501-115-16	Sequence 16, Appl
21	299.5	17.0	150	4 US-09-501-115-24	Sequence 24, Appl
22	299.5	17.0	172	4 US-09-501-115-46	Sequence 46, Appl
23	297.5	16.9	465	4 US-09-252-9191A-18919	Sequence 18919, Appl
24	296	16.8	155	4 US-09-501-115-48	Sequence 48, Appl
25	273.5	15.6	375	4 US-09-134-000C-5347	Sequence 5347, Appl
26	270	15.4	206	4 US-09-540-236-3768	Sequence 3768, Appl
27	265	15.1	346	4 US-09-489-0394-9885	Sequence 9885, Appl

Run on:	July 6, 2004, 13:34:55 ;	Search time 13.8278 Seconds
		(without alignments)
		1217.121 Million cell updates/sec
Title:	US-09-857-518A-31	
Perfect score:	1758	
Sequence:	1 ETGATAVRKFVLYCGVCHSD.....ERRVKKDVRFRREVIDVENTL 326	
Scoring table:	BLOSUM62	
Gapop:	10.0 , Gapext 0.5	
Searched:	389414 seqs, 5165971 residues	
Total number of hits satisfying chosen parameters:	389414	
Minimum DB seq length: 0		
Maximum DB seq length: 2000000000		
Post-processing: Maximum Match 0%		
Listing first 45 summaries		
Database :	Issued_Patents_AA:*	
	1: /cgn2_6_ptodata/2/iaa/5A_COMB.pep:*	
	2: /cgn2_6_ptodata/2/iaa/5B_COMB.pep:*	
	3: /cgn2_6_ptodata/2/iaa/5A_COMB.pep:*	
	4: /cgn2_6_ptodata/2/iaa/5B_COMB.pep:*	
	5: /cgn2_6_ptodata/2/iaa/PCOTUS_COMB.pep:*	
	6: /cgn2_6_ptodata/2/iaa/backfiles1.pep:*	
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.		
SUMMARIES		
RESULT 1		
	US-09-501-115-12	
	; Sequence 12, Application US/09521115	
	; Patent No. 6552249	
	; GENERAL INFORMATION:	
	; APPLICANT: Cahoon, Rebecca E.	
	; ATTORNEY: Pader, Gary M.	
	; TITLE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs	
	; FILE REFERENCE: B1120 US NA	
	; CURRENT APPLICATION NUMBER: US/09/501,115	
	; CURRENT FILING DATE: 2000-02-09	
	; EARLIER APPLICATION NUMBER: 60/119,585	
	; EARLIER FILING DATE: 1999-February-10	
	; NUMBER OF SEQ ID NOS: 48	
	; SOFTWARE: Microsoft Office 97	
	; SEQ ID NO: 12	
	; LENGTH: 355	
	; TYPE: PRT	
	; ORGANISM: Glycine max	
	US-09-501-115-12	
	Query Match	57.3%
	Best Local Similarity	59.6%
	Matches 201; Conservative 39; Mismatches 24; Indels 5;	
	QY	1 ETGATDVREKVLKPLRPTCENCITHLENYCPNLQIOTYGSKYDGTMTYGGSNMNTDBHFIYRI 59
	Db	28 ETGEKDLVPLKQXGIGCHSDHMKRNEDWTGPIVGHVYEVGCKVKKPK-SNR 59
	QY	60 QGRGMWLHGRPLRPTCENCITHLENYCPNLQIOTYGSKYDGTMTYGGSNMNTDBHFIYRI 119
	Db	88 VGVGMIGSCKR-SCEBSCDENLYCPNLQIOTYGSYDGTMTYGGSNMNTDBHFIYRI 146
	QY	120 PDNEPLDGAPlLCAgGTTTSPWYYGLKPGMELJGVEMVPRPFSRSPPNLPGINGSRQLQ 179
	Db	147 PDNEPLDAAAPLCAgGTTTSPWYYGLDKPGLNLGV-----VCGGGL-GHMAV 194
	QY	180 SLVPPLIKEGGSYCgTSPLMHSIgT-----DODOMEAMSTMgIDTIVPAVRP 229
	Db	195 KFAKALGANVTVTISPNKCREAENIGADSFVVSREQDQMAgNGTMgIDTIVSAPR 254
	QY	230 LEPLISLTKNGVTVGTAQVPLDLPPVPLIGRKYAGVSAIGGMKETOEMIDFAAEHN 289
	Db	255 LVPJGLKPKHGLVAMVGAPEKPLJELPPTESLMGRMgGSSICgMKTQEMIDFAAKG 314
	QY	290 ITADIEVLPIDYNTAMERVKCDVRFREVIDVENTL 326
	Db	315 VKEPLIEVLPIDYNTAERLAKADVKYAFVIDGNTL 351

ALIGNMENTS

RESULT 2

US-09-501-115-4

Sequence 4, Application US/09501115

Patent No. 6552249

GENERAL INFORMATION:

APPLICANT: Cahoon, Rebecca E.

APPLICANT: Fader, Gary M.

APPLICANT: Rafalski, Antoni

TITLE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs

FILE REFERENCE: BB1328 US NA

CURRENT APPLICATION NUMBER: US/09/501,115

CURRENT FILING DATE: 2000-02-09

EARLIER APPLICATION NUMBER: 60/119,585

EARLIER FILING DATE: 1999-February-10

NUMBER OF SEQ ID NOS: 48

SOFTWARE: Microsoft Office 97

SEQ ID NO 4

LENGTH: 360

TYPE: PRT

ORGANISM: Glycine max

US-09-501-115-4

Query Match 47.3%; Score 832; DB 4; Pred. No. 1.2e-81;

Best Local Similarity 48.8%; Mismatches 84; Indels 40; Gaps 6;

Matches 168; Conservative 52; Gaps 6;

Qy 2 TGATDVRKVLGRPLPCTNCIHLHENYCPNLIGTYGKVSQDTHMANKDQWSTYPIVPGHLYVNTDEHFTVRI 119

Db 40 TGDDDVYVTKLILYCGHSQDLSHSKNDWKNRKYAQPMPGHEAGEYEVGVQVTKPAGDRY 99

Qy 61 G-RCWLGRLRPLPCTNCIHLHENYCPNLIGTYGKVSQDTHMANKDQWSTYPIVPGHLYVNTDEHFTVRI 119

Db 100 GVGCMVNS-CQSCS-CDKFENHCPGMWTFYNSVDRGTRTHGGSYSSMVVHESFVWRF 157

Qy 120 PDNLPLDGAAPLCLGAGTTSPWRYGLDAPGMHIVENRFRSRPMLPGLWGSRLQ 179

Db 158 PDAMPDQKGAFLCLGAGTTSPWRYGLDAPGMHIVENRFRSRPMLPGLWGSRLQ 179

Qy 180 SLVPPPLIKEGGSYGTSPALMHS-----LLRTDODMERAAMSTMDGID 222

Db 202 ---HYAWKFGRAFGNKTIVSSSPGKQEAUERLGADAFVSKOADEMRATMSNDGII 258

Qy 223 TYPATPPLPEPLSILKTKGKVUTVYJAVQPLDLPKPLPLGKMYTAGSAIGGMKETQEMI 282

Db 259 TVSANYPMAPLFGKLPKPGNRMIVGVLPEXIEVPEALVARNKTLAGSCTGMRDTCQEMI 318

Qy 283 DFAABENITADLEPIVLDINTAMBRVKCDVRFREVIDVNTL 326

Db 319 DLAARKHGVTADIEVIGAEYNTAMERLAKADVRYREVIDVNTL 362

RESULT 4

US-09-501-115-8

Sequence 8, Application US/09501115

Patent No. 6552249

GENERAL INFORMATION:

APPLICANT: Cahoon, Rebecca E.

APPLICANT: Fader, Gary M.

APPLICANT: Rafalski, Antoni

FILE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs

CURRENT APPLICATION NUMBER: US/09/501,115

CURRENT FILING DATE: 2000-02-09

EARLIER APPLICATION NUMBER: 60/119,585

EARLIER FILING DATE: 1999-February-10

NUMBER OF SEQ ID NOS: 48

SOFTWARE: Microsoft Office 97

SEQ ID NO 8

LENGTH: 361

TYPE: PRT

ORGANISM: Zea mays

NAME/KEY: UNSURE

LOCATION: (3.9)

FEATURE:

NAME/KEY: UNSURE

LOCATION: (78)

US-09-501-115-8

Query Match 45.5%; Score 800; DB 4; Pred. No. 3.5e-78;

Best Local Similarity 47.7%; Mismatches 83; Indels 46; Gaps 7;

Matches 165; Conservative 52; Gaps 7;

Qy 2 TGATDVRKVLGRPLPCTNCIHLHENYCPNLIGTYGKVSQDTHMANKDQWSTYPIVPGHLYVNTDEHFTVRI 115

Db 33 TGDDDVYVTKLILYCGHSQDLSHSKNDWKNRKYAQPMPGHEAGEYEVGVQVTKPAGDRY 92

Qy 61 G-RCWLGRLRPLPCTNCIHLHENYCPNLIGTYGKVSQDTHMANKDQWSTYPIVPGHLYVNTDEHFTVRI 119

Db 93 GVGCMVNS-CQSCS-CDKFENHCPGMWTFYNSVDRGTRTHGGSYSSMVVHESFVWRF 150

Qy 120 PDNLPLDGAAPLCLGAGTTSPWRYGLDAPGMHIVENRFRSRPMLPGLWGSRLQ 173

Db 151 PDAMPDQKGAFLCLGAGTTSPWRYGLDAPGMHIVENRFRSRPMLPGLWGSRLQ 199

RESULT 3

US-09-501-115-14

Sequence 14, Application US/09501115

Patent No. 6552249

GENERAL INFORMATION:

APPLICANT: Cahoon, Rebecca E.

APPLICANT: Fader, Gary M.

APPLICANT: Rafalski, Antoni

TITLE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs

FILE REFERENCE: BB1328 US NA

CURRENT APPLICATION NUMBER: US/09/501,115

CURRENT FILING DATE: 2000-02-09

EARLIER APPLICATION NUMBER: 60/119,585

EARLIER FILING DATE: 1999-February-10

NUMBER OF SEQ ID NOS: 48

SOFTWARE: Microsoft Office 97

SEQ ID NO 14

LENGTH: 311

TYPE: PRT

ORGANISM: Triticum aestivum

US-09-501-115-14

RESULT 5
 US-09-501-115-10
 ; Sequence 10, Application US/09501115
 ; Patent No. 655249
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; FILER: Fader, Gary M.
 ; TITLE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs
 ; FILE REFERENCE: BB1328 US NA
 ; CURRENT FILING DATE: 2010-02-09
 ; EARLIER APPLICATION NUMBER: 60/119,585
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 2
 ; LENGTH: 364
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa
 US-09-501-115-2

Query Match 41.7%; Score 732.5; DB 4; Length 364;
 Best Local Similarity 43.8%; Pred. No. 7.6e-71; Mismatches 55; Indels 59; Gaps 5;
 Matches 152; Conservative 55;

Qy 6 DVFREKVLYCGVCHSDIIMAKNDGTSPIVPGHELYGVVTBVKCKVKKFKS----W 58
 Db 34 DVIKTKVLYCGICLTDLHVKNDDGMAFPVPGHEIVTVGAGTGYKAKDTVGVY 93
 Qy 59 RGCRCWLHGRLRPCTCENCTHLENCPMLCITYGSKYD---GMMTCGGSNNMNTDHF 115
 Db 94 FVASC-----RGCECCGNGYNYCARMVTTGNGVHDGGRATGGSFSDALTWNHRY 146
 Qy 116 IVEIDPFDGAPLLCAGITIYSPWYVGLKPGMGLGVENPRPFRSPRPLNIPGDI- 173
 Db 147 VLRVPAHGLPDSAPLICAGVTTYSERVTHGLNAPGKGV-----VGLGIGH 195
 Qy 174 -----WGSRLQSLVPLPLIKEGSGSTGTSP----ALMH-----SLRSTDQDOMEAMS 215
 Db 196 VAKKFAKFGARTV-----STSPGKQEAELHGADEFDVSRDAGMAAA 244
 Qy 216 TMGDTIITDTPAVRPLEPLISLTKTNGKVYTVGIAVOPDLPVFLIGRKWVAGSAIGGM 275
 Db 245 TMGDLNTVSAWPIAPDSLMKPMQNVFGGPTRPLSLPAVAVPGGKGITGNCVGJ 304

Qy 276 KETQEMIDFAERNTIADEIIVIDYLTAMERVVKODYRERVIDY 322
 Db 305 RDCQAMLDFAGERGCTAEEVIMDYNTAMERLLEKNDVYRERVIDY 351

RESULT 7
 US-09-501-115-6
 ; Sequence 6, Application US/09501115
 ; Patent No. 655249
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; FILER: Fader, Gary M.
 ; TITLE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs
 ; FILE REFERENCE: BB1328 US NA
 ; CURRENT FILING DATE: 2010-02-09
 ; EARLIER APPLICATION NUMBER: 60/119,585
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 6
 ; LENGTH: 358
 ; TYPE: PRT
 ; ORGANISM: Triticum aestivum
 US-09-501-115-6

Query Match 40.8%; Score 717; DB 4; Length 358;
 Best Local Similarity 43.1%; Pred. No. 3.6e-69; Mismatches 93; Indels 40; Gaps 6;
 Matches 147; Conservative 61;

Qy 6 DVFREKVLYCGVCHSDIIMAKNDGTSPIVPGHELYGVVTBVKCKVKKFKS----SW 58
 Db 33 DVIKTKVLYCGICLTDLHVKNDDGMAFPVPGHEIVTVGAGTGYKAKDTVGVY 92

59 RQGRCWLHGRLPRCNCIHLHENYCPNLQTYGSKYD---GTMIGGYSNNMVTDEH 114
 93 FDSC-----RSCDSCSKYYESCPQVETNSGYSLDDGGATTIKGFFSDALVHQR 145
 115 FIVTRPDNLPLDGAAPLLCAGITTYSPMRYGLDKPGMHGVENPRRSPRSPNPLGGL- 173
 146 YVVRVEASLPLAGAAPLLCAGVTFSPMVQGLAPKHLGV-----vSLGGLG 194
 174 -----WGSRLQSLVPLPPLIKEGGSKYGTSPALHSLRLTDQDMEAAMSTMGLIDTV 224
 195 HLAVERGKAPKQVTVTISGKRDPEQDMRAAGTLDQVDTV 253
 225 FAVRPLEPLSLLKTKNGKVVTVGIAVQPLDPVPLIGKMWASSAIGKRETOEMIDE 284
 254 SADHPVVPVPLDILKEMQMVGLPLKPLQYPAFSLVAGKCRVASSAGGGYECQAMLDF 313
 285 AEHNTTADIEVPIPDYLNTAMERYVKDTRFRFVTDVENT 325
 314 AGERGHTADVEVGMQDVNTAQLRDLNDYRFRFVDDASS 354

RESULT 8
 US-09-501-115-32
 ; Sequence 32, Application US/09501115
 ; Patent No. 6552249
 ; GENERAL INFORMATION:
 ; APPLICANT: Cahoon, Rebecca E.
 ; APPLICANT: Fader, Gary M.
 ; APPLICANT: Rafalski, Antoni
 ; TITLE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs
 ; FILE REFERENCE: BBI328 US NA
 ; CURRENT APPLICATION NUMBER: US/09/501,115
 ; CURRENT FILING DATE: 2000-02-09
 ; EARLIER APPLICATION NUMBER: 60/119,585
 ; EARLIER FILING DATE: 1999-February-10
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO: 30
 ; LENGTH: 358
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (201)
 US-09-501-115-30

Query Match 36.9%; Score 648; DB 4; Length 358;
 Best Local Similarity 41.0%; Pred. No. 1.1e-61;
 Matches 141; Conservative 58; Mismatches 103; Indels 42; Gaps 9;
 Query 2 TGATDVRFKVLYCGVCHSDTHMAKNDWGTSYTYPIVPGHELGVVTEVGCKVKKPSWROG 61
 Db 32 TGPDDVYTKHYCGICHSDHQIKNDLGMNSYPMYFGHEVGELEVGSNSYSPF--RVG 88
 Query 62 RCWLGRLRPTCEN---C1HLENCPNLLQTYGSKYDGMUPTGCGYNNMVTDEHIVR 118
 Db 89 ELVGVLGLVGCKNQCCQDIENTSKWISYNDVYDQKPTQGSPETMVEQKVK 148
 Query 119 IPDNPLDGAAPLCAAGITTYSPWRYGLDKPGMHGVENPRRSPRSPNPLGGL-WGS 176
 Db 149 IPEGIAPQVAPLICAGTVTSPLYFGKLGKESLGRGGI-----LGLGGVGHMGV 197
 Query 177 RLQSLVPLKPLKEGSYGT-----S2ALMH----SILRTDQDMEAMSTDGIIID 222
 Db 198 KIAAXL-----GHVTVVSSSDKKKQEALEHLGADQYLVSSDATAQERADSUYIID 250
 Query Match 37.8%; Score 665; DB 4; Length 360;
 Best Local Similarity 40.6%; Pred. No. 1.6e-63;
 Matches 141; Conservative 57; Mismatches 97; Indels 52; Gaps 7;
 Query 1 ETGATDVRFKVLYCGVCHSDTHMAKNDWGTSYTYPIVPGHELGVWTEVGCKVKEFKSWR- 59
 Db 32 KTGPEVDVVLKVKGICHTDVKNDLGA SKYFNVPGHEVGEVTEVGDEVSKEPAGDV 91
 Query 60 -----QGRCWLHGRLPRCNCIHLHENYCPNLQTYGSKYDGTMTYGGYSNNMVTDE 113
 Db 92 VGVGTVGCC-----RDRPCKRANEQVQNKWISYNDVTDKPTQGFSAMVVDQ 144
 Query 114 HEIVRPDNLPLDGAAPLCAAGITTYSPWRYGLDKPGMHGVENPRRSPRSPNPLGGL 173
 Db 145 KFVVKIPAGAEPQAPLICAGTVTSPLYFGKLGKESLGRGGI-----LGLGGV 193
 Query 174 WGSRLQSLVPLPPLIKEGGSKYGTSPALMES-----LLRTDQDMEAMST 216
 Db 194 GHMG-----VKYAKSMGHHTVISSSNKRAAMDDLGADATIVSSDTQAAAADS 245
 Query 217 MDGIDTVPAVRPLEPLISLTKNGKVVTVGIAVQPLDPVFLI-TGRNVAAGSAIGGM 275
 Db 246 LDYIIDTVPKHPLEPLVALLRMDGKLVNGVIAELSE-VSPMWMGLRKLTGFIGSM 304
 Query 276 KETOIDEAHEHTTADIEVPIPDYLNTAMERYVKDTRFVTDVENT 322
 Db 305 DETEIVVQFCVTDKGTLTSQIVVVKMDYVNOAFERLERSNDVYRFVVDV 351

RESULT 10
 US-09-328-352-7473
 ; Sequence 7473, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary J. Breton et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO: 7473
 ; LENGTH: 353
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 ; US-09-328-352-7473

Query Match 35.9%; Score 630.5; DB 4; Length 353;
 Best Local Similarity 39.9%; Pred. No. 8.6e-60;

Matches	139;	Conservative	52;	Mismatches	94;	Indels	63;	Gaps	9;	Qy	284	FAAEHNITADIEVTPIDYINTAMERVVERDVRFRFVIDV	322
Qy	6	DYDPRKLYCGVCSIDINAKNDIGTSPVNPCHELVWVNTGKCKKKR-----	56	Db	365	FCRANGIACDIEMTEIQINQTERMILSEDVTRFVIDM	403						
Db	34	DVYIQIEVGVCHSDLHQAKNDIGFSAVPLVCPHEIVGRTS1GPKVTKYKGDLVIGC	93										
Qy	57	---SWRQGRCLWHLGRRLRPTCENCHHLENY-C-PNLQIOTG-S-KYDGTMTYGGYSANMVT	111		RESULT 12								
Db	94	MVDSCR-----TCSAHSGLQCEPGNQTYGGDTRDPTGYSQSCTC	141	US-09-501-115-20									
Db	112	DEHFIVRIPDNLPLDGAAPLCLAGTTSPWYGLDKPGMELGVWPRFRSPRPLNL	171	Sequence 20	Application US/09501115								
Qy	142	SQDEVLKPVENLDAQAVPLPLCAGITWSPLRHWNVGK-GSKVAVF-----	186	Patent No.	6552249								
Qy	172	GLMSRQSIQVPLVPLIKEGGSYGTSPRNL-----HSLRITDQDOMEAMS	215	GENERAL INFORMATION:									
Db	187	GLG-----LGEWAALKLNLGAEVLTTSANKEQDAKQLAHHVVLSTDENQMKSYLN	241	APPLICANT: Cabon, Rebecca E.									
Qy	216	TMDGIIIDVPAVPLPLEPLISLILKINGVTVVGAVQPLDLPVPLITGRKWAAGSATGGM	275	APPLICANT: Fader, Gary M.									
Db	242	QFLIVIVTPYHDLKPYPTPLAINGTVWVGV-LGE-TANSUPMILGRKSTAGSIVG	300	APPLICANT: Ralski, Antoni									
Qy	276	KTFQEMIDFAAEHNITADIEVTPIDYINTAMERVVKDVRFRFVIDV	323	TITLE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs									
Db	301	RETQELLDCEGHNIVSEVEMINONINEAERMLKSDVYKRFVIDM	348	FILE REFERENCE: BBL329 US NA									
				CURRENT APPLICATION NUMBER: US/09/501,115									
				CURRENT FILING DATE: 2000-02-39									
				EARLIER APPLICATION NUMBER: 60/119,585									
				EARLIER FILING DATE: 1999-February-10									
				NUMBER OF SEQ ID NOS: 48									
				SEQ ID NO: 20									
				TYPE: PRT									
				ORGANISM: Glycine max									
				US-09-501-115-20									
				Query Match Score 32.9%; Best Local Similarity 40.1%; Matches 139; Conserv	32.9%	Score 578;	DB 4;						
				ative 44; Indels 116; Gaps 7;		DB 54;	Length 357;						
Qy	93	VGTGTVNSCRDCEYCNQDQEYH-----TRGSVYTFNGVDFDSTTIGGSSYIVHERYC	147	Qy	3	GATDVRFKVLYCGVCHSDIHMADNGTSTYPTVPHLGWVTEVGCKVTKPFSWRQGR	62						
Db	93	-----AVKFGKAFLGSLVYTFESTSISKKEAISLGLADKFVUSSNQEMTALAKSLDF	248	Db	33	GNEDVH1KITHCGVCFADWVTRKHGD5KSKYD-----YOGMTYGGYSNNMMYTDHF1	116						
Qy	117	VRIDPNLFDGAPLCLAGTTSPWYGLDKPGMHLGVWPRFRSPRSPPLNLPGLGWS	176	Qy	63	CWLHGRRLPCTCENCNTHLENYCPNLLCITYGSKY-----YOGMTYGGYSNNMMYTDHF1	116						
Db	148	FMIKPSYPLASIAEPLCAGITVSPMVRKMNQGSKSLGV-----IGUGGLGHM	196	Db	93	VGTGTVNSCRDCEYCNQDQEYH-----TRGSVYTFNGVDFDSTTIGGSSYIVHERYC	147						
Qy	177	RLQSLVPLPKE3SYGTSPALM-----SLRLTD-----QDOMEAAMS TMDG	219	Db	117	VRIDPNLFDGAPLCLAGTTSPWYGLDKPGMHLGVWPRFRSPRSPPLNLPGLGWS	176						
Db	197	-----AVKFGKAFLGSLVYTFESTSISKKEAISLGLADKFVUSSNQEMTALAKSLDF	248	Qy	117	VRIDPNLFDGAPLCLAGTTSPWYGLDKPGMHLGVWPRFRSPRSPPLNLPGLGWS	176						
Qy	220	LIIDTVPAVPRPLSPLSLIKTKNGKVYTVGIAVQPLDLPVPLITIGRKMVAGSAIGCMKETQ	279	Db	148	FMIKPSYPLASIAEPLCAGITVSPMVRKMNQGSKSLGV-----IGUGGLGHM	196						
Db	249	-----AVKFGKAFLGSLVYTFESTSISKKEAISLGLADKFVUSSNQEMTALAKSLDF	307	Qy	93	-----SLRLTD-----QDOMEAAMS TMDG	219						
Qy	64	WLGRLRPLPCTCENCNTHLENYC--PHLQIOTG-SKXYYDGMTYGGYSNNMMYTDHF1	116	US-09-489-030A-10-08									
Db	143	VGTCGCMVTDSCRHCAACBSCBIEQCAEYPTL-----TYNQDRDHLGPTFGSERTVVSARFV	200	Sequence 10-08, Application US/0948903A									
Qy	117	VRPDLNLPDGAAPLCLAGTTSPWYGLDKPGM-----HLGVNPRPFSR	165	Patent No.	661036								
Db	201	LRTQGLDPLTAAPLCLAGTTSPWYGLDKPGM-----HLGVNPRPFSR	259	GENERAL INFORMATION:									
Qy	166	PPNLPLGWLGSRLQSLLPPLIKEGGSYGTSPALMHSLLRTDODOMEAMSTMDG1LIDTVP	225	APPLICANT: Gary Breton et. al.									
Db	260	-----VALFTRSPGSEEARRLGADQV-----ILSTDPRQMAVAGRFDLIDTVP	305	TITLE OF INVENTION: PNEUMONIA FOR DIAGNOSTICS AND THERAPEUTICS									
Qy	226	AVRPLEPLISLILKINGVTVTGIAVQPLDLPV-----PLIIGRKWVAGSAIGCMKETQEMD	283	CURRENT APPLICATION NUMBER: US/09/489,039A									
Db	306	HQDHLNPVTLAKRDTGTVLWGL-LEPLBPAVHGQLOMGRMRRSTAGS1GIAETQEMD	364	CURRENT FILING DATE: 2000-01-27									
				PRIOR APPLICATION NUMBER: US 60/117,747									
				PRIOR FILING DATE: 1999-01-29									
				NUMBER OF SEQ ID NOS: 14342									
				SEQ ID NO: 10508									
				LENGTH: 358									
				TYPE: PRT									

ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10508

Query Match 32.3%; Score 567.5; DB 4; Length 358;
Best Local Similarity 38.3%; Pred. No. 6e-53;
Matches 133; Conservative 56; Mismatches 105; Indels 53; Gaps 10;

Qy 1 BTGADTVRKVLYCGVCHSDHMAKNDWGTSVTPGHEVNGVTVBVGCKVKRE-----55
Db 33 DPGQDQVIAAYCGVCHSDIHOARSEWAGTLYCPGHEVGRITAVGNAVSRVAVGDL 92

Qy 56 -----KSWRQGRCWLGRLPCTCNCIHLLENTPNLIQTYGSKYDGT-MTYGGYN 107
Db 93 VGVGCRVDSOO-----CEBCABELENTCDHMTVITNGPTQDAGDHTUGGYSQ 140

Qy 108 NMVTDDEHFTVIRP-DNLPLDGAAPLCAAGITTSVPRYYGLDKPGMHLGVWPRRERSRP 166
Db 141 QIVVNEYRVLRIITHPBAQLAAPLCAAGITTSVPRYYGLDKPGMHLGVWPRRERSRP 189

Qy 167 PLNLDCG-----WSRQLSLLVPPLIKEGGSYGSPALMHS-----LLR7DQDMEAMAS 215
Db 190 -VGIGLGKNGIKLAAHAMGAHVV--AFTTSESKENAARALGADVVVSNEDEAAHVK 245

Qy 216 TMDGIIIDTVAPRLEPLISLKINGVQPLDIP-VFPLIIGRKVAGSAIGG 274
Db 246 SEDFLINTVAPAPHNLDAAFTLKLKGDTMILVGAPATPHPSPEVNLIFRRSIASSMIGG 305

Qy 275 MKETOEMIDPAEAHENTTADIRVPIPDYLTAMERVYKCDYVTRFRFVTD 321
Db 306 IPETQEMLDPAEAHGVIVADIELRGDEINEAWERAVKGDVTRFRFVTD 352

RESULT 14
US-09-501-115-28

Sequence 28; Application US/09501115
Patient No. 655249

GENERAL INFORMATION:
APPLICANT: Cahoon, Rebecca E.
APPLICANT: Rafalski, Antoni
TITLE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs
CURRENT APPLICATION NUMBER: US/09/501,115
EARLIER APPLICATION NUMBER: US1328735 NA
CURRENT FILING DATE: 2000-02-09
EARLIER FILING DATE: 1999-February-10
NUMBER OF SEQ ID NOS: 48
SOFTWARE: Microsoft Office 97
SEQ ID NO 28
LENGTH: 360

TYPE: PRT
ORGANISM: Glycine max
US-09-501-115-34

Query Match 21.6%; Score 379; DB 4; Length 360;
Best Local Similarity 29.4%; Pred. No. 1.8e-32; Mismatches 55; Indels 62; Gaps 10;
Matches 103; Conservative 55; Mismatches 130; Indels 62; Gaps 10;

Qy 3 GATDVRKVLYCGVCHSDIHLKNDWGTSVTPGHELYGVWTRVCKYFKFKSWRQGR 62
Db 32 GDFDVDIKILACGVCVCSASDLETVGGNGEKEPLCVCYHEINGEARVPGKTVLIRGQRVG 91

Qy 63 CWLHGRLRPTCNCIHLLENTPNLIQTYGSKYD-GTMVYGGSYNNMMTDEHFIVRID 121
Db 92 VGAQSYSCLOCRQCNNDNETCILHOLDITYGAWWTGTIVSQGQTASHVRTHEHWYFPID 151

Qy 122 NPLDQGAPLCAPILAGITTSVPRYVOLGDKPGMHLGVWPRFSSRSPNLLGGLGSRQLSL 181
Db 152 GLKTERAAPLCAPIGATTAISPLVRNGCG_PSKKVG-----V3LGGI-----191

Qy 182 VPPLIKEGGSYGTS-----PALMSLLR7DQDMEAMASMDGIIIDTVPA-----VRPL 230
Db 192 -----GHFQGMFAKALGAETTAISRTAKEADAKLQGADGFILATADKDNKHEHIMPF 243

Qy 231 EPLIS-----LLKTKNGKTYTYVGIAV--QPLDLPVFPFLITGRKXWAGSAI 272
Db 244 DLIVNTASSSGFNSEYELLDVHGWVSVGLPGEQGQID--NFQLRNGVLGGSH 301

Qy 273 GGMKETQEMIDPAEAHNTTADIEVPI- DYLNTAERVYKDVTRFRFV 320
Db 302 GSREKVLAMLDLAVETKGINSWETDISEGKLTALTRHNNDVTRFV 351

Query Match 31.8%; Score 559; DB 4; Length 292;
Best Local Similarity 41.6%; Pred. No. 3.7e-52;
Matches 124; Conservative 44; Mismatches 88; Indels 42; Gaps 9;

Qy 1 EITGATDVRKVLYCGVCHSDIHLKNDWGTSVTPGHELYGVWTRVCKYFKFKSWR 59
Db 16 ENGDNDVTIKLXCGICHTDLYHARNNEWGTTMYPVPGHEIGVWTRKOVERDVKCFKEGDR 75

Qy 60 QG-RCWLHGRLRPTCNCIHLLENTPNLIQTYGSKYD-GTMVYGGSYNNMMTDEHFIVR 118
Db 76 VGVGCLAAASCLE--CEHCKTDXDENYCEKLFQVNGVFWGSDITGGYSQIFVADYRYVH 133

Qy 119 IPONLPLDGAAPLCAAGITTSVPRYYGL_DKPGMHLGVWPRRERSRSPNLLPGLGSR 177
Db 134 IPENLAMDAAAPLCAAGITVNPFLKDHLYVAPSKEKIGV-----VGLGG- 177

Search completed: July 6, 2004, 13:40:16
Job time : 14.8278 secs

Scoring table:	BLOSUM62	Gapcp:	10.0	Gapext:	0.5
title:	US-09-857-518A-31	length:	0	maximum DB seq length:	2000000000
perfect score:	1758	number of hits satisfying chosen parameters:	1276540	post-processing:	Minimum Match 0% Maximum Match 100% Listing first 45 summaries
sequence:	1 ETGATDVRKFVLYCGVCHSD.....	searched:	1276540 seqs, 311283816 residues	database :	Published Applications AA.*
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	1: /cgm2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*	RESULT 1 US-10-424-599-266181
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	2: /cgm2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*	; Sequence 266181, Application US/10424599
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	3: /cgm2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*	; Publication No. US2004031072A1
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	4: /cgm2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*	; GENERAL INFORMATION:
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	5: /cgm2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*	; APPLICANT: La Rosa Thomas J
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	6: /cgm2_6/ptodata/2/pubpa/US08_PUBCOMB.pep:*	; APPLICANT: Kovalic David K
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	7: /cgm2_6/ptodata/2/pubpa/US08_PUBCOMB.pep:*	; APPLICANT: Zhou Yihua
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	8: /cgm2_6/ptodata/2/pubpa/US10_PUBCOMB.pep:*	; APPLICANT: Cao Yongwei
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	9: /cgm2_6/ptodata/2/pubpa/US10_PUBCOMB.pep:*	; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	10: /cgm2_6/ptodata/2/pubpa/US09_PUBCOMB.pep:*	; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	11: /cgm2_6/ptodata/2/pubpa/US09C_PUBCOMB.pep:*	; FILE REFERENCE: 38-21532231B
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	12: /cgm2_6/ptodata/2/pubpa/US10A_PUBCOMB.pep:*	; CURRENT APPLICATION NUMBER: US/10/424,599
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	13: /cgm2_6/ptodata/2/pubpa/US10B_PUBCOMB.pep:*	; NUMBER OF SEQ ID NOS: 285684
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	14: /cgm2_6/ptodata/2/pubpa/US11C_PUBCOMB.pep:*	; SEQ ID NO: 266181
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	15: /cgm2_6/ptodata/2/pubpa/US11A_PUBCOMB.pep:*	; LENGTH: 359
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	16: /cgm2_6/ptodata/2/pubpa/US11_PUBCOMB.pep:*	; OTHER INFORMATION: Glycine max
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	17: /cgm2_6/ptodata/2/pubpa/US60_PUBCOMB.pep:*	; NUMBER OF SEQ ID NOS: 285684
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	18: /cgm2_6/ptodata/2/pubpa/US60_PUBCOMB.pep:*	; LENGTH: 359
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	SUMMARIES
result No.	Score	Query Match	Length	DB ID	Description
1	1008	57.3	359	12 US-10-424-599-266181	Sequence 266181,
2	1007	57.3	355	12 US-10-424-599-12	Sequence 261443,
3	1007	57.3	362	9 US-09-947-150-2	Sequence 2, Appli
4	979	55.7	362	9 US-09-947-150-2	Sequence 2, Appli
5	979	55.7	362	13 US-10-091-009-2	Sequence 2, Appli
6	979	55.7	362	13 US-10-091-009-2	Sequence 2, Appli
7	958	54.5	361	12 US-10-424-599-266973	Sequence 266973,
8	939	52.9	360	12 US-10-425-114-5801	Sequence 152603,
9	884	50.3	365	12 US-10-424-599-152603	Sequence 356617,
10	884	50.3	365	12 US-10-425-114-36617	Sequence 54709,
11	884	50.3	365	12 US-10-424-599-152604	Sequence 152604,
12	880	50.1	366	12 US-10-425-114-44445	Sequence 44445,
13	880	50.1	366	12 US-10-425-114-44445	Sequence 44445, A
14	882	49.4	360	12 US-10-424-599-205468	Sequence 205468,
15	882	49.4	364	12 US-10-424-599-205468	Sequence 205468,
16	842	47.9	371	14 US-10-425-114-5786-14	Sequence 14, Appli
17	832	47.3	381	14 US-10-425-114-64190	Sequence 6, Appli
18	810	46.1	381	14 US-10-425-114-64190	Sequence 10, Appli
19	800	45.5	361	14 US-10-425-114-886-8	Sequence 118472,
20	790	44.9	366	14 US-10-425-114-886-8	Sequence 204767,
21	779	44.3	428	16 US-10-425-114-965-119472	Sequence 39429, A
22	755.5	43.0	360	16 US-10-425-114-965-119472	Sequence 2, Appli
23	741	42.2	378	12 US-10-425-114-39429	Sequence 10, Appli
24	732.5	41.7	364	14 US-10-357-886-6	Sequence 12, Appli
25	725	41.2	370	14 US-10-361-450-10	Sequence 56149, A
26	725	41.2	406	12 US-10-425-114-56149	Sequence 40577, A
27	724.5	41.2	373	12 US-10-425-114-40577	Sequence 45843, A
28	718.5	40.9	370	12 US-10-425-114-4583	Sequence 62384, A
29	718.5	40.9	370	12 US-10-425-114-62284	Sequence 6, Appli
30	717	40.8	358	14 US-10-357-886-6	Sequence 12, Appli
31	716	40.7	358	14 US-10-361-450-12	Sequence 59052, A
32	716	40.7	378	12 US-10-425-114-59052	Sequence 39395, A
33	716	40.7	390	12 US-10-425-114-39395	Sequence 43324, A
34	716	40.7	390	12 US-10-425-114-39519	Sequence 39519, A
35	716	40.7	391	12 US-10-425-114-71567	Sequence 71567, A
36	714.5	40.6	379	12 US-10-437-96-17457	Sequence 172457,
37	713.5	40.6	336	16 US-10-437-96-17457	Sequence 123886,
38	712.5	40.5	391	16 US-10-437-96-17457	Sequence 11, Appli
39	709	40.3	359	14 US-10-361-450-11	Sequence 176237,
40	707	40.2	410	16 US-10-37-96-176237	Sequence 54812, A
41	697.5	39.7	329	12 US-10-424-54812	Sequence 230622,
42	696	39.6	384	12 US-10-424-54812	Sequence 230622,
43	679	38.6	357	9 US-09-947-027-9	Sequence 9, Appli
44	679	38.6	357	13 US-10-091-099-9	Sequence 13927, A
45	670.5	34.1	344	15 US-10-369-493-13927	
Scoring table:		Alignments:			ALIGNMENTS
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	1: ETGATDVRKFVLYCGVCHSD.....	RESULT 1 US-10-424-599-266181
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	2: /cgm2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*	; Sequence 266181, Application US/10424599
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	3: /cgm2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*	; Publication No. US2004031072A1
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	4: /cgm2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*	; GENERAL INFORMATION:
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	5: /cgm2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*	; APPLICANT: La Rosa Thomas J
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	6: /cgm2_6/ptodata/2/pubpa/US07_PUBCOMB.pep:*	; APPLICANT: Kovalic David K
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	7: /cgm2_6/ptodata/2/pubpa/US08_PUBCOMB.pep:*	; APPLICANT: Zhou Yihua
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	8: /cgm2_6/ptodata/2/pubpa/US08_PUBCOMB.pep:*	; APPLICANT: Cao Yongwei
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	9: /cgm2_6/ptodata/2/pubpa/US09_PUBCOMB.pep:*	; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	10: /cgm2_6/ptodata/2/pubpa/US09_PUBCOMB.pep:*	; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	11: /cgm2_6/ptodata/2/pubpa/US09C_PUBCOMB.pep:*	; FILE REFERENCE: 38-21532231B
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	12: /cgm2_6/ptodata/2/pubpa/US10A_PUBCOMB.pep:*	; CURRENT APPLICATION NUMBER: US/10/424,599
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	13: /cgm2_6/ptodata/2/pubpa/US10B_PUBCOMB.pep:*	; NUMBER OF SEQ ID NOS: 285684
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	14: /cgm2_6/ptodata/2/pubpa/US11C_PUBCOMB.pep:*	; SEQ ID NO: 266181
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	15: /cgm2_6/ptodata/2/pubpa/US11A_PUBCOMB.pep:*	; LENGTH: 359
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	16: /cgm2_6/ptodata/2/pubpa/US11_PUBCOMB.pep:*	; OTHER INFORMATION: Glycine max
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	minimum DB seq length:	0	17: /cgm2_6/ptodata/2/pubpa/US60_PUBCOMB.pep:*	; NUMBER OF SEQ ID NOS: 285684
run on:	July 6, 2004, 13:38:51 ; Search time 36.036 Seconds (without alignments), 2816.033 Million cell updates/sec	maximum DB seq length:	0	18: /cgm2_6/ptodata/2/pubpa/US60_PUBCOMB.pep:*	; LENGTH: 359
Scoring table:		Organism:			ORGANISM: Glycine max
Scoring table:		Feature:			FEATURE: PRT
Scoring table:		Other information:			OTHER INFORMATION: Glycine max
Scoring table:		Clone ID:			CLONE ID: PAT_MRT3847_82383C.1.pep
Scoring table:		SRQ ID:			SRQ ID NO: 266181
Scoring table:		Length:			US-10-424-599-266181
Scoring table:		Query Match			Query Match
Scoring table:		Length:			Length
Scoring table:		DB ID			DB ID
Scoring table:		Description			Description

No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

IMMATURES

Documentation
Presentation
Query
Mutant
Transcript
Protein
%
Result

Query	Subject	Score	Length	Start	End	Mismatches	Indels	Gaps	7:
Query match	29.3%	Score 1087	Length 352						
Best Local Similarity	59.5%	Score 1127	Length 352						
Matches	203; Conservative	6.3e-97	Pred. No. 6	42; Mismatches	64; Indels	32; Gaps	7;		
Qy	1 ERGATDVRKFVLYTCGVCHSDTHMAKNDWSTSTYPIVPGHBLVGVTTVEGVCKVKRKEPK-SWR	59	59						
Db	3.4 EPGEKDVAFRVLGRLRPTENCHLENVCPNLIIQTYGSKYDGTMITYGGYSMMVTDHEFTVRI	93	93						
Qy	6.0 QGRWLRHLGRPLRPTENCHLENVCPNLIIQTYGSKYDGTMITYGGYSMMVTDHEFTVRI	119	119						
Db	9.4 VGVGCLVIDSSCR-TCQNCCDNLENVCPQSTFTYGAKYRDTITYGGYSDSMTADEHFVVR	152	152						
Qy	12.0 PDNLPLDGAAPLCAAGITTYSPRYYGCDKPGMHLGVENFRFRFSRPPNLPGWGSRLQ	179	179						
Db	15.3 PDRPLPLDAAAPLCAAGITTYSPRYYGCDKPGMHLGVENFRFRFSRPPNLPGWGSRLQ	196	196						

RESULT 3
 US-10-424-599-261443 12 ; Sequence 12, Application US/10357886
 ; Publication No. US20030159170A1
 ; GENERAL INFORMATION: Cahnor, Rebecca B.
 ; APPLICANT: Fader, Gary M.
 ; APPLICANT: Rafalski, Antoni
 ; TITLE OF INVENTION: Plant Cannamyl-Alcohol Dehydrogenase Homologs
 ; FILE REFERENCE: B1128 US NA
 ; CURRENT APPLICATION NUMBER: US/10/357,886
 ; CURRENT FILING DATE: 2003-02-04
 ; PRIOR APPLICATION NUMBER: US/09/501,115
 ; PRIOR FILING DATE: 2000-02-09
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/119,585
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-February-10
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO: 12
 ; LENGTH: 355
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; SEQ ID NO: 261443
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE: OTHER INFORMATION: Clone ID: PNT_MRT3847_78105C.1.pep
 ; US-10-424-599-261443

Query Match 57.3%; Score 1007; DB 12; Length 362;
 Best Local Similarity 59.6%; Pred. No. 8.2e-97;
 Matches 201; Conservative 39; Mismatches 73; Indels 24; Gaps 5;

Qy 1 ETGATDVREKYLICGCVHSDIHMAKNDWGTSTPIVPGHELWGVVTEVGCKVKKEF-SWR 59
 Db 35 ETGEKOLVPEKVKYQCGCHSDIHMAKNDWGTSTPIVPGHELWGVVTEVGCKVKKEFQKVGD 94

Qy 60 QGRCLWHLGRPLTCENYCPNLIQTTGSKYVDTGMMYGGSNMVTDEEFIVTRI 119
 Db 95 VGVGMGCSRSCESCDENLYNCPRMILTYGVKYDGTITHGGSIDLMDAEEFIVTRI 153

Qy 120 PDNLPLDGAAPLCAAGITTSFWRYCGLDKPGNHLGVWPRFRSRPPLNGLWGSRLQ 179
 Db 154 PDNLPLDAAAPLCAAGITVSLRYYGLDKPGNHLGV-----VGLGGL-GHMAV 201

Qy 180 SLVPPLIKEGGSGTSPALMHSILRT-----DODQMRAMSTMIDGILDITDYPAVRP 229
 Db 202 KPAKALGANVTVTISNKKKEAATGADSPVSVSREQDQMRAMGMGMDGIDTVSATVHP 261

Qy 230 LEPLISLILKNGKVTVYGLAVQPLDLPVPLIGRKNVAGSAGIGMMKETQEMDFAEHN 289
 Db 262 LVPLIGLKPAAKLYVGAPEPPLPEVFSLIMGRKNVGGSSIGGMKETQEMDFAAHHG 321

Qy 290 ITADIEVYPIDVTANTAMERYKDVRFVTDVENTL 326
 Db 322 VPKDIEVYPIDVTANTERLAKADVXTRFVTDIGNTL 358

RESULT 4
 US-10-557-886-12
 ; Sequence 12, Application US/10357886
 ; Publication No. US20030159170A1
 ; GENERAL INFORMATION: Cahnor, Rebecca B.
 ; APPLICANT: Fader, Gary M.
 ; APPLICANT: Rafalski, Antoni
 ; TITLE OF INVENTION: Plant Cannamyl-Alcohol Dehydrogenase Homologs
 ; FILE REFERENCE: B1128 US NA
 ; CURRENT APPLICATION NUMBER: US/10/357,886
 ; CURRENT FILING DATE: 2003-02-04
 ; PRIOR APPLICATION NUMBER: US/09/501,115
 ; PRIOR FILING DATE: 2000-02-09
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/119,585
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-February-10
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO: 12
 ; LENGTH: 355
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; SEQ ID NO: 261443
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE: OTHER INFORMATION: Clone ID: PNT_MRT3847_78105C.1.pep
 ; US-10-424-599-261443

Query Match 57.3%; Score 1007; DB 14; Length 355;
 Best Local Similarity 59.6%; Pred. No. 7.9e-97;
 Matches 201; Conservative 39; Mismatches 73; Indels 24; Gaps 5;

Qy 1 ETGATDVREKYLICGCVHSDIHMAKNDWGTSTPIVPGHELWGVVTEVGCKVKKEF-SWR 59
 Db 28 ETGZKDLVFKVQICGCHSDIHMAKNDWGTSTPIVPGHELWGVVTEVGCKVKKEF-SWR 87

Qy 60 QGRCLWHLGRPLTCENYCPNLIQTTGSKYVDTGMMYGGSNMVTDEEFIVTRI 119
 Db 88 VGVGMGCSRSCESCDENLYCPMILTYGVYDGTITHGGSIDLMDAEEFIVTRI 146

Qy 120 PDNLPLDGAAPLCAAGITTSFWRYCGLDKPGNHLGVWPRFRSRPPLNGLWGSRLQ 179
 Db 147 PDNLPLDAAAPLCAAGITVSLRYYGLDKPGNHLGV-----VGLGGL-GHMAV 194

Qy 180 SLVPPLIKEGGSGTSPALMHSILRT-----DODQMRAMSTMIDGILDITDYPAVRP 229
 Db 195 KPAKALGANVTVTISNKKKEAATGADSPVSVSREQDQMRAMGMGMDGIDTVSATVHP 254

Qy 230 LEPLISLILKNGKVTVYGLAVQPLDLPVPLIGRKNVAGSAGIGMMKETQEMDFAEHN 289
 Db 255 LVPLIGLKPAAKLYVGAPEPPLPEVFSLIMGRKNVGGSSIGGMKETQEMDFAAHHG 314

Qy 290 ITADIEVYPIDVTANTAMERYKDVRFVTDVENTL 326
 Db 315 VPKDIEVYPIDVTANTERLAKADVXTRFVTDIGNTL 351

RESULT 5
 US-10-424-599-261443 13 ; Sequence 13, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION: APPICNT: T. Ross, et al.
 ; APPLICANT: T. Ross, et al.
 ; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 ; FILE REFERENCE: 38-21/532231B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO: 261443
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE: OTHER INFORMATION: Clone ID: PNT_MRT3847_78105C.1.pep
 ; US-10-424-599-261443

Query Match 55.7%; Score 979; DB 9; Length 362;
 Best Local Similarity 56.6%; Pred. No. 7.2e-94;
 Matches 196; Conservative 39; Mismatches 67; Indels 44; Gaps 6;

2 TOATDVFKEKULVGSVCHSDIHMAKNDWGTSTPIVPGHELWGVVTEVGCKVKKEF-SWR-
 Qy 2 TOATDVFKEKULVGSVCHSDIHMAKNDWGTSTPIVPGHELWGVVTEVGCKVKKEF-SWR-

Db 35 TGEEDVRFKVLTCGICHSIDLHSIKNDWGFMSMPLYVGHEIVGEETTEVGSKVVKVNVGDKV 94
 Qy 61 -----GRCHLHGRRLPCTCNCIHLLENCTNLQTYGSKYXKGDMTYGGYSNNMVTDEH 114
 95 GVGCLVGAC-----HSCECSCANDLENCYCPMLTYASYTHDGTITYGSKYSDHANVNER 147
 Db 115 FIVRFDNLPLDGAAPLCCGTTSPWRYGLDKPFGMHLGSMWPRFRSPRSPNLPGLW 174
 Qy 148 YITRFDDNMPDGGAPLCCGTTSPWRYGLDKPFGMHLGSMWPRFRSPRSPNLPGLW 174
 Db 148 YITRFDDNMPDGGAPLCCGTTSPWRYGLDKPFGMHLGSMWPRFRSPRSPNLPGLW 193
 Qy 175 GSRLQSLVPPPLKEGSS---YGTSPALMHSLRT-----DQDMQAAMSTMDGJ 220
 Db 194 G-----LGHVAVKFAKAGSKVTVTISTSSKKEALKNGADSFLVSEDBQBNQAAAGTLDGJ 251
 Db 221 IDTVPAVRFPLISLKTNGKVTVGAYQPLDLPVPEPLIGRMVAGSAIGGKETOE 280
 Qy 252 IDTVPAVRFPLISLKTNGKVTVGAYQPLDLPVPEPLIGRMVAGSAIGGKETOE 311
 Db 281 MIDEAAHNITADIEVPIDYLNTAMERYVKDVRFRFVIDVENTL 326
 Qy 312 MIDEAAHNITADIEVPIDYLNTAMERYVKDVRFRFVIDVENTL 357

RESULT 5
 US-09-947-027-2
 ; Sequence 2, Application US/09947027
 ; Patent No. US20020124281A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chiang, Vincent Lee C.
 ; INVENTOR: Li, Laigeng
 ; TITLE OF INVENTION: METHODS FOR SIMULTANEOUS CONTROL OF LIGNIN CONTENT AND COMPOSITION IN PLANTS
 ; FILE REFERENCE: 066040-9718
 ; CURRENT APPLICATION NUMBER: US/09/947,027
 ; CURRENT FILING DATE: 2001-09-05
 ; PRIOR APPLICATION NUMBER: 60/230,086
 ; PRIOR FILING DATE: 2000-09-05
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO: 2
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: aspen populus tremuloides

RESULT 6
 US-10-091-009-2
 ; Sequence 2, Application US/10091009
 ; Sequence 2, Application US/10091009
 ; Publication No. US20020138870A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chiang, Vincent Lee C.
 ; INVENTOR: Li, Laigeng
 ; TITLE OF INVENTION: METHODS FOR SIMULTANEOUS CONTROL OF LIGNIN CONTENT AND COMPOSITION IN PLANTS
 ; FILE REFERENCE: 066040-9718
 ; TITLE OF INVENTION: CELLULOSE CONTENT IN PLANTS
 ; CURRENT APPLICATION NUMBER: US/10/091,009
 ; CURRENT FILING DATE: 2002-03-06
 ; PRIOR APPLICATION NUMBER: 09/947,027
 ; PRIOR FILING DATE: 2001-09-05
 ; PRIOR APPLICATION NUMBER: 60/230,086
 ; PRIOR FILING DATE: 2000-09-05
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: Patentin version 3.0
 ; SEQ ID NO: 2
 ; LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: aspen populus tremuloides

Query Match 55.7%; Score 979; DB 13; Length 362;
 Best Local Similarity 56.6%; Pred. No. 7.e-94;
 Matches 39; Mismatches 67; Indels 44; Gaps 6;

Query 2 TGTATVRFKVLTCGICHSIDLHMAKDGTTSPYPLVPEGLVGTWVYGCVKVKEKSWRQ-60
 Db 35 TGEEDVRFKVLTCGICHSIDLHMAKDGTTSPYPLVPEGLVGTWVYGCVKVKEKSWRQ-60
 Qy 61 -----GRCHLHGRRLPCTCNCIHLLENCTNLQTYGSKYDGMTYGGYSNNMVTDEH 114
 Db 95 GVGCLVGAC-----HSCECSCANDLENCYCPMLTYASYTHDGTITYGSKYSDHANVNER 147
 Qy 115 FIVRFDNLPLDGAAPLCCGTTSPWRYGLDKPFGMHLGSMWPRFRSPRSPNLPGLW 174
 Db 148 YITRFDDNMPDGGAPLCCGTTSPWRYGLDKPFGMHLGSMWPRFRSPRSPNLPGLW 193
 Qy 175 GSRLQSLVPPPLKEGSS---YGTSPALMHSLRT-----DQDMQAAMSTMDGJ 220
 Db 194 G-----LGHVAVKFAKAGSKVTVTISTSSKKEALKNGADSFLVSEDBQBNQAAAGTLDGJ 251
 Qy 221 IDTVPAVRFPLISLKTNGKVTVGAYQPLDLPVPEPLIGRMVAGSAIGGKETOE 280
 Db 252 IDTVSAVRFPLISLKTNGKVTVGAYQPLDLPVPEPLIGRMVAGSAIGGKETOE 311
 Qy 281 MIDEAAHNITADIEVPIDYLNTAMERYVKDVRFRFVIDVENTL 326
 Db 312 MIDEAAHNITADIEVPIDYLNTAMERYVKDVRFRFVIDVENTL 357

RESULT 7
 US-10-424-599-266973
 ; Sequence 266973, Application US/10424599
 ; Publication No. US200403107A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J
 ; INVENTOR: Kovacic, David K
 ; INVENTOR: Zhou, Yihua
 ; FILE REFERENCE: 38-211532231B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285634
 ; SEQ ID NO: 266973
 ; LENGTH: 361
 ; TYPE: PRT

Query Match 55.7%; Score 979; DB 9; Length 362;
 Best Local Similarity 56.6%; Pred. No. 7.e-94;
 Matches 39; Mismatches 67; Indels 44; Gaps 6;

Query 2 TGTATVRFKVLTCGICHSIDLHMAKDGTTSPYPLVPEGLVGTWVYGCVKVKEKSWRQ-60
 Db 35 TGEEDVRFKVLTCGICHSIDLHMAKDGTTSPYPLVPEGLVGTWVYGCVKVKEKSWRQ-60
 Qy 61 -----GRCHLHGRRLPCTCNCIHLLENCTNLQTYGSKYDGMTYGGYSNNMVTDEH 114
 Db 95 GVGCLVGAC-----HSCECSCANDLENCYCPMLTYASYTHDGTITYGSKYSDHANVNER 147
 Qy 115 FIVRFDNLPLDGAAPLCCGTTSPWRYGLDKPFGMHLGSMWPRFRSPRSPNLPGLW 174
 Db 148 YITRFDDNMPDGGAPLCCGTTSPWRYGLDKPFGMHLGSMWPRFRSPRSPNLPGLW 193
 Qy 175 GSRLQSLVPPPLKEGSS---YGTSPALMHSLRT-----DQDMQAAMSTMDGJ 220
 Db 194 G-----LGHVAVKFAKAGSKVTVTISTSSKKEALKNGADSFLVSEDBQBNQAAAGTLDGJ 251
 Qy 221 IDTVPAVRFPLISLKTNGKVTVGAYQPLDLPVPEPLIGRMVAGSAIGGKETOE 280
 Db 252 IDTVSAVRFPLISLKTNGKVTVGAYQPLDLPVPEPLIGRMVAGSAIGGKETOE 311
 Qy 281 MIDEAAHNITADIEVPIDYLNTAMERYVKDVRFRFVIDVENTL 326
 Db 312 MIDEAAHNITADIEVPIDYLNTAMERYVKDVRFRFVIDVENTL 357

Organism: Glycine max
 Feature:
 OTHER INFORMATION: Clone ID: PAT_MRT3847_830388C.1.pep
 JS-10-424-599-266973

Query Match 54.5%; Score 958; DB 12; Length 361;
 Best Local Similarity 56.6%; Pred. No. 1.2e-91;
 Matches 194; Conservative 43; Mismatches 70; Indels 36; Gaps 6;

1 BTGATDVRPKVLYCGVCHSDIIMAKNDWGTSTYPIVPGHLIGVNTTEVGCKVCKFKPSWRQ 63
 34 KTGKEDVTPKVLYCGVCHSDIIMKLKNEWSDSTYPIVPGRLIGVGETTEVGSKVDFKVGDK 93
 61 -----GRCWLRGRDRPCTCNCIHLHENYCPNLIQTYGSKYDSEMTMTGYSNNMVTDE 113
 94 VAAGCLVGSC-----HSQCNVNLLNENYCQVQVPTYGRKYDGTGTTGFSDEMAVE 146
 114 HPIVRIPDNLPLDGAAPLICAGLITYSPWXYGLDKPGMELGVVPPRRFRSRPPLNPLG 173
 147 HFWVNIPSALEDDAAPLICAGLITYPGLDKPGMELGV-----VGLGL 195

174 WGSRLQSLVLPPLIKEKGSYCTSP-----SLURTDODMBAAMSTMMDGIIIDT 223
 196 -GHIAVKFAZGLKLVTVISTSPKCKNEA1QHLAGDFFYVSRDQMQAMCTLDGIIIDT 254

224 VPAVPLPLEPLSLLKTKNGKTVGIAVQPLDLPVPLIGRKMAGSATGGMKETQEMID 283
 255 VSAVEPLMLPLDKSHGKLVAVGAAPEKEPLLLPPLIGRKSLAGSYIGKIKETQEMIN 314

284 FAAEINNTADAEVIPDYINTAMERYVKDVRFREVIDENTL 326
 315 FAAEINVRPELEVPMDDYNTAMERLQAKDVKYREVIDGNTL 357

RESULT 9
 US-10-424-599-152603
 Sequence 152603, Application US/10424599
 Publication No. US2004031072A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J
 APPLICANT: Kovalic, David K
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 FILE REFERENCE: 38-21153223; B
 CURRENT APPLICATION NUMBER: US/10/424, 599
 CURRENT FILING DATE: 2003-04-28
 SEQ ID NO: 152603
 LENGTH: 356
 TYPE: PRT
 ORGANISM: Glycine max
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT3847_108825C.1.pep
 US-10-424-599-152603

Query Match Score 884; DB 12; Length 356;
 Best Local Similarity 50.6%; Pred. No. 6.9e-84;
 Matches 173; Conservative 49; Mismatches 84; Indels 36; Gaps 3;

1 BTGATDVRPKVLYCGVCHSDIIMAKNDWGTSTYPIVPGHLIGVNTTEVGCKVCKFKPSWRQ 60
 30 ENGVDDVTLKLECQWCHSILHETLKDNGTTPVPGHLIGVNTTEVGCKVCKFKPSWRQ 89

61 GRCMLHGRLRPTCNCIHLHENYCPNLIQTYGSKYDGMITYGYSNNMVTDEHFLVRIPI 120
 90 VGVGVIVESKECVCQDQDNLGAGLTCAGTIVSPMKYGMTEGKHLGV-----AGLGG---- 149

121 DNLEDDGAAPLICAGLITISPVRYGLDKPGMHLGVWPPFRSRPPLNPLGLWMSRLQS 160
 150 ENLPDAGAPLICAGTIVSPMKYGMTEGKHLGV-----AGLGG---- 190

181 LVPPHKEGGSGYGTSPALMHS-----LIRTDQDQMEAMSTMIDGIIIDT 223
 191 LGTPRIKLAAPGKLVTVTSSPKNQAEADRLGADFLYFLSSDPAKMKALGTDYIDT 250

224 VPAVPLPLEPLISLKTNGKTVGIAVQPLDLPVPLIGRKMAGSATGGMKETQEMID 283
 251 ISAVESLPIHGLKXNGLTVGLPNKPLBLPFLPVAGRKLGGSNFSLKETQEMID 310

284 FAAEINNTADAEVIPDYINTAMERYVKDVRFREVIDENTL 325
 311 FCGHENNTADIELTQMDDQINTAMERLSRADVKYFVVIDYASS 352

RESULT 10
 US-10-424-599-152617
 Sequence 386517, Application US/10425114
 Publication No. US2004034888A1
 GENERAL INFORMATION:
 APPLICANT: Liu, Jingdong

Organism: Glycine max
 Feature:
 OTHER INFORMATION: Clone ID: UG-GMROPI033A03_FLI.pep
 JS-10-425-114-55801

Query Match 52.9%; Score 929.5; DB 12; Length 360;
 Best Local Similarity 56.2%; Pred. No. 1.1e-88;
 Matches 190; Conservative 40; Mismatches 73; Indels 35; Gaps 6;

1 BTGATDVRPKVLYCGVCHSDIIMAKNDWGTSTYPIVPGHLIGVNTTEVGCKVCKFKPSWRQ 60
 42 KTGKEDVTPKVLYCGVCHSDIIMKLKNEWSDSTYPIVPGRLIGVGETTEVGSKVDFKVGDK 101

61 -----GRCWLRGRDRPCTCNCIHLHENYCPNLIQTYGSKYDGMITYGYSNNMVTDE 113
 102 VAAGCLVGSC-----HSQCNVNLLNENYCQVQVPTYGRKYDGTGTTGFSDEMAVE 154
 114 HPIVRIPDNLPLDGAAPLICAGITYSPIVPGHLIGVWPPFRSRPPLNPLG 173

1 APPLICANT: Zhou, Yihua
 1 APPLICANT: Kovacic, David K.
 1 APPLICANT: Screen, Steven E.
 1 APPLICANT: Tabaska, Jack E.
 1 APPLICANT: Cao, Yongwei
 1 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 1 Title of Invention: Plants and Uses Thereof for Plant Improvement
 1 FILE REFERENCE: 38-21 (53313) B
 1 CURRENT FILING DATE: 2003-04-28
 1 SEQ ID NO: 36617
 1 LENGTH: 65
 1 TYPE: PRT
 1 ORGANISM: Glycine max
 1 FEATURE:
 1 OTHER INFORMATION: Clone ID: LIB3040-049-D9_FLI.pep
 US-10-425-114-36617

Query Match Score 884; DB 12; Length 365;
 Best Local Similarity 50.3%; Pred. No. 7.1e-84;
 Matches 173; Conservative 49; Mismatches 84; Indels 36; Gaps 3;

Qy 1 ETGATDVRKVLYCGVCHSDIHMAKNDGTSTYPIVPGHLYGVYTCYKVKFMSRQ 60
 Db 39 ENYDDVTILKILFGVCHSDLHTLKDNGTTPVPGHLYGVYTCYNNVKFGDK 98

Qy 61 GRCWLHGRLRPTCNCIHLHENYCPNLCTYGSKYDGTMTYGGYSNNMTDEHFTYRIP 120
 Db 99 VGVGVIVESCKECCNCCQDLENCPVPGHLYGVYTCYNNVKFMSRQ 158

Qy 121 DNPLDGAPLLCAAGTTTSPWYGLDKPQGMHGVNPRFRSRPPLNPGMSRLQS 180
 Db 159 ENPLDAGPPLCAAGTTTSPWYGLDKPQGMHGVNPRFRSRPPLNPGMSRLQS 325

Qy 181 LVPPLIKEGGSYGTSPALMHS-----
 Db 200 LGHYAIKLAFAKGLKVTWISSSPNQAEADRLGADFFLSSDPAKMKAALGTMIDYIITD 283

Qy 224 VPAYRPLEPLISLTKNGKVTWISSSPNQAEADRLGADFFLSSDPAKMKAALGTMIDYIITD 319

Qy 284 FAEEHNITADIEVLPIDYNTAMERVVKDVRFFIVDVENT 325
 Db 320 FCGHNITADIEVLPIDYNTAMERVVKDVRFFIVDVENT 325

RESULT 12
 US-10-424-599-152604
 ; Sequence 152604, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J.
 ; APPLICANT: Kovalic David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: SCY Nucleic Acid Molecules and Other Molecules Associated With
 ; Title of Invention: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21 (53323) B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO: 152604
 ; LENGTH: 360
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_108826C.1.pep
 US-10-424-599-152604

Query Match Score 880; DB 12; Length 360;
 Best Local Similarity 50.1%; Pred. No. 1.8e-83;
 Matches 173; Conservative 49; Mismatches 84; Indels 36; Gaps 3;

Qy 1 ETGATDVRKVLYCGVCHSDIHMAKNDGTSTYPIVPGHLYGVYTCYNNVKFGDK 60
 Db 34 ENYDDVTILKILFGVCHSDLHTLKDNGTTPVPGHLYGVYTCYNNVKFGDK 93

Qy 61 GRCWLHGRLRPTCNCIHLHENYCPNLCTYGSKYDGTMTYGGYSNNMTDEHFTYRIP 120
 Db 94 VGVGVIVESCKECCNCCQDLENCPVPGHLYGVYTCYNNVKFMSRQ 180

Qy 181 LVPPLIKEGGSYGTSPALMHS-----
 Db 154 ENPLDAGPPLCAAGTTTSPWYGLDKPQGMHGVNPRFRSRPPLNPGMSRLQS 194

Qy 224 VPAYRPLEPLISLTKNGKVTWISSSPNQAEADRLGADFFLSSDPAKMKAALGTMIDYIITD 283

RESULT 11
 US-10-425-114-54709
 ; Sequence 54709, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovacic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; Title of Invention: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21 (53313) B
 ; CURRENT FILING DATE: 2003-04-28
 ; SEQ ID NO: 73128
 ; LENGTH: 365
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: JC-GMFL02220073B02_FLI.pep
 US-10-425-114-54709

Qy 121 DNPLDGAPLLCAAGTTTSPWYGLDKPQGMHGVNPRFRSRPPLNPGMSRLQS 223

RESULT 13

Db 255 ISAVSILIPILGLKLKNGKLTVGLPNKPELPFLPVAGRLIGESNFGIKETQEMLD 314 ; CURRENT FILING DATE: 2003-02-04 ; PRIOR APPLICATION NUMBER: US/09/501.115

Qy 284 PAAEINNITADLEVIPIPDLYNTAMERVVKDVRFRFIDVENT 325 ; PRIOR FILING DATE: 2003-02-09 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/119,585

Db 315 FCAAEINNITADLELIKMDQINTAMERLSKADVKYRFIDVANS 356 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-February-10 ; NUMBER OF SEQ ID NOS: 48 ; SOFTWARE: Microsoft Office 97 ; SEQ ID NO: 4

RESULT 13

Db US-10-425-114-4445 ; Sequence 4445, Application US/10425114 ; GENERAL INFORMATION: ; APPLICANT: Zhou, Yihua ; APPLICANT: Kovacic, David K. ; APPLICANT: Screen, Steven E. ; APPLICANT: Tabaska, Jack E. ; APPLICANT: Cao, Yongwei ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with FILE REFERENCE: 38-21(53.313)B ; CURRENT APPLICATION NUMBER: US/10/425,114 ; CURRENT FILING DATE: 2003-04-28 ; NUMBER OF SEQ ID NOS: 73128 ; SEQ ID NO: 44445 ; LENGTH: 366 ; TYPE: PRT ; ORGANISM: Glycine max ; FEATURE: OTHER INFORMATION: Clone ID: 700954121_FU1_pep US-10-425-114-4445

Query Match 50.1%; Score 880; DB 12; Length 366; Best Local Similarity 50.6%; Pred. No. 1.9e-83; Mismatches 49; Indels 36; Gaps 3; Matches 173; Conservative 37; APPLICANT: Zhou, Yihua ; APPLICANT: Kovacic, David K. ; APPLICANT: Screen, Steven E. ; APPLICANT: Tabaska, Jack E. ; APPLICANT: Cao, Yongwei ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with FILE REFERENCE: 38-21(53.313)B ; CURRENT APPLICATION NUMBER: US/10/425,114 ; CURRENT FILING DATE: 2003-04-28 ; NUMBER OF SEQ ID NOS: 73128 ; SEQ ID NO: 44445 ; LENGTH: 366 ; TYPE: PRT ; ORGANISM: Glycine max ; FEATURE: OTHER INFORMATION: Clone ID: 700954121_FU1_pep US-10-425-114-4445

Query Match 50.1%; Score 880; DB 12; Length 366; Best Local Similarity 50.6%; Pred. No. 1.9e-83; Mismatches 49; Indels 36; Gaps 3; Matches 173; Conservative 37; APPLICANT: Zhou, Yihua ; APPLICANT: Kovacic, David K. ; APPLICANT: Screen, Steven E. ; APPLICANT: Tabaska, Jack E. ; APPLICANT: Cao, Yongwei ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with FILE REFERENCE: 38-21(53.313)B ; CURRENT APPLICATION NUMBER: US/10/425,114 ; CURRENT FILING DATE: 2003-04-28 ; NUMBER OF SEQ ID NOS: 73128 ; SEQ ID NO: 44445 ; LENGTH: 366 ; TYPE: PRT ; ORGANISM: Glycine max ; FEATURE: OTHER INFORMATION: Clone ID: 700954121_FU1_pep US-10-425-114-4445

RESULT 14

Db US-10-357-886-4 ; Sequence 4, Application US/10357886 ; GENERAL INFORMATION: ; APPLICANT: Cahoon, Rebecca E. ; APPLICANT: Fader, Gary M. ; APPLICANT: Rafalski, Antoni ; TITLE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs ; FILE REFERENCE: BB1328 US NA ; CURRENT APPLICATION NUMBER: US/10/357,886

Qy 284 PAAEINNITADLEVIPIPDLYNTAMERVVKDVRFRFIDVENT 325 ; PRIOR FILING DATE: 2003-02-09 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/119,585

Db 321 FCAAEINNITADLELIKMDQINTAMERLSKADVKYRFIDVANS 362 ; NUMBER OF SEQ ID NOS: 48 ; SOFTWARE: Microsoft Office 97 ; SEQ ID NO: 4

RESULT 15

Db US-10-424-599-205468 ; Sequence 205468, Application US/10424599 ; GENERAL INFORMATION: ; APPLICANT: La Rosa, Thomas J. ; APPLICANT: Kovacic, David K. ; APPLICANT: Zhou, Yihua ; APPLICANT: Cao, Yongwei ; TITLE OF INVENTION: Saccharomyces Cerevisiae Nucleic Acid Molecules and Other Molecules Associated with Plants and Uses Thereof for Plant Improvement ; FILE REFERENCE: 38-21(53.223)B ; CURRENT APPLICATION NUMBER: US/10/424,599 ; CURRENT FILING DATE: 2003-04-28 ; NUMBER OF SEQ ID NOS: 285644 ; SEQ ID NO: 205468 ; LENGTH: 364 ; TYPE: PRT ; ORGANISM: Glycine max ; FEATURE: OTHER INFORMATION: Clone ID: PAT_MRT3847_27566C.1.pep US-10-424-599-205468

Query Match 47.9%; Score 842; DB 12; Length 364; Best Local Similarity 49.6%; Pred. No. 1.9e-79; Mismatches 50; Indels 44; Gaps 6; Matches 171; Conservative 50; APPLICANT: Cahoon, Rebecca E. ; APPLICANT: Fader, Gary M. ; APPLICANT: Rafalski, Antoni ; TITLE OF INVENTION: Plant Cinnamyl-Alcohol Dehydrogenase Homologs ; FILE REFERENCE: BB1328 US NA ; CURRENT APPLICATION NUMBER: US/10/357,886

Qy 3 GATDVRFKVLYCGVCHSDTHIMAKNDWGTSTTYPVPGHELVGVVTFVGCKYFKF-----56 ; PRIOR FILING DATE: 2003-02-09 ; PRIOR APPLICATION NUMBER: 60/119,585

Db 38 GDNDTFLDVLGIGCHIDLAKVRNDRFISIYPMVPGHEIVGKVTKVRGRTVKFSVNGDIAG 97 ; NUMBER OF SEQ ID NOS: 48 ; SOFTWARE: Microsoft Office 97 ; SEQ ID NO: 4

57 -SWRCGRCPMLHGRLRPTCENICHHLBNCYCNLICITYGSKYDGMITYGGYSNNRNYTDEHF 115
 Db 98 VGGSVGSC-----GSCSGCSNGFVYCKMMLTYSARYHYDGTTITQGTSNNTVVDQNF 150
 Qy 116 IVRIPDNLPDGAAPLCAAGITTSPPRYGLDKPGMHLGCVENPRFRSRPPLNLPGLNG 175
 Db 151 VVLIPKSLPDLGAALLCAAGITTSPPKCYGLAQPSLHGV-----VGLGS 196
 Qy 176 SRLQSLVPLIPLIKESSY---GTSPLMHS-----LLRPTDQHRAAMSSTMDGTI 221
 Db 197 -LGHVAVKPKAKAFRMHVTWTSSEKKEALEXKGADEBLVSLDOOQUDARGTMDAII 254
 Qy 222 DTVPAVRPLEPLISLILKTKKVVITGLAVOPLDLPPFPLTIGRKVAGSAIGMKEQTM 281
 Db 255 DTVSANHSQCPPLIAALKTSCKLILNGGPUSPLEVLAAMPILGRMOMIAGSAGGPERETQW 314
 Qy 282 IDFAAEHNITADIEVLPIDYINTAMERYVTKDVRERFVIDVENTL 326
 Db 315 MDFAAKHNITADVEVPMDDVNTAERLENDVXRFVIDVANTL 359

Search completed: July 6, 2004, 13:48:42
 Job time : 37.036 secs

SCOP protein - protein search, using sw model														
Run on:	July 6, 2004, 13:29:59	Search time	8.38046 Seconds											
Sequence:	1 ETGATDVRFKVLVCGVHSQ.....	(without alignments)												
Scoring table:	BLOSUM62			2025.529 Million cell updates/sec										
Title:	US-09-857-518A-31			ERVKRDVRFRFVIDVENTL 326										
Perfect score:	1759													
Score:	1													
Searched:	141681 seqs, 52070155 residues			141681										
Total number of hits satisfying chosen parameters:														
Minimum DB seq length: 0														
Maximum DB seq length: 2000000000														
Post-processing: Minimum Match 0%														
Maximum Match 100%														
Listing first 45 summaries														
Database :	SwissProt_42;*													
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.														
SUMMARIES														
Result No.	Score	Query Match	Length	DB ID	Description									
1	1092	62.1	359	1 MTD_FRAAN	Q9zrf1 fragaria an									
2	971	55.2	361	1 MTD_MEDSCR	P93271 meembryant									
3	963	54.8	337	1 MTD_BTICR	P42734 petroselinum									
4	960	54.6	357	1 MTD_ARATH	Q02971 arabidopsis									
5	959	54.6	359	1 MTD2_ARATH	Q02972 arabidopsis									
6	945	53.8	365	1 MTD_APIGR	Q38707 apium grave									
7	884	50.3	360	1 MTDH_ARATH	P42734 arabidopsis									
8	870.5	49.5	359	1 MTD_MEDSA	Q81250 medicago sa									
9	860.5	48.9	363	1 MTD3_STYTHU	Q43118 stylosanthe									
10	833.5	47.4	354	1 MTD1_STYHU	Q43137 stylosanthe									
11	740	42.1	357	1 CAD2_PICAB	Q82035 picea abies									
12	740	42.1	357	1 CAD2_PICAB	Q82030 picea abies									
13	739	42.3	357	1 CADH_PINTA	P41637 Pinus taeda									
14	723	41.2	357	1 CADH_PINTA	Q40976 pinus radiata									
15	691.5	39.3	360	1 CAD1_ARACO	P42120 arachis hyp									
16	686	39.3	357	1 CAD9_TOBAC	P30360 nicotiana t									
17	681	38.7	357	1 CAD4_TOBAC	P31659 nicotiana t									
18	666	37.9	357	1 CADH_PORDE	P31657 populus del									
19	656	37.3	361	1 CADH_DOLPR	O22380 lolium pere									
20	654	37.2	356	1 CADH_EUCGU	P31657 eucalyptus									
21	653	37.1	356	1 CADH_EUCGL	O64956 eucalyptus									
22	646	36.7	357	1 CAD2_ARATH	O49462 arabidopsis									
23	641	36.5	358	1 CADH_MEDSA	P31656 medicago sa									
24	629	35.8	354	1 CAD1_EUCGU	P31657 eucalyptus									
25	628	35.7	367	1 CADH_MAIZE	O24562 zea mays (m									
26	621.5	35.4	355	1 CADH_EUCBO	P50746 eucalyptus									
27	621	35.3	365	1 CAD1_ARATH	P48523 arabidopsis									
28	611	34.8	365	1 CADH_SACOF	O82036 saccharum o									
29	599.5	34.1	346	1 CADH_WYCTU	P31915 mycobacteriu									
30	555.5	31.6	349	1 YAHK_ECOLI	P75591 escherichia									
31	379	21.6	360	1 YW97_BEAST	Q04894 saacharomy									
32	365.2	20.8	339	1 YJGB_ECOLI	P72320 escherichia									
33	356.5	20.3	339	1 ADH3_BACST	P42128 bacillus st									

34	350.5	19.9	337	1	ADH1_BACST	P12311	bacillus st	
35	350.5	19.9	339	1	ADH2_BACST	P42327	bacillus st	
	320.5	18.2	361	1	YCZ5_YEAST	P25377	saccharomyces	
36	316.5	18.0	340	1	ADH4_YHME	O1186	rhizobium m	
37	338	300.5	352	1	ADH3_EMENTI	P47754		
	339	294.5	348	1	ADH2_KLUMA	P49383	kluyveromyces	
	40	285.5	16.2	348	1	ADH2_KRNL	O94038	candida alb
	41	278	15.8	349	1	ADH1_ASPPFL	P41785	aspergillus
	42	277.5	15.8	350	1	ADH1_CANAL	P43067	candida alb
	43	277.5	15.8	353	1	ADH1_NEUCR	O9p6c8	neurospora
	44	275.5	15.7	344	1	ADH1_KLULIA	P20369	kluyveromyces
	45	269.5	15.3	347	1	ADH2_KLUMA	O9p4c2	kluyveromyces

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 Searched: 141681 seqs, 52070155 residues
 Total number of hits satisfying chosen parameters: 141681
 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum March 0%

Database : SwissProt_42:*

SUMMARIES						
Result	Query	Match	Length	DB	ID	Description
No.	Score					
1	1092	62.1	359	1	MTD_FRAAN	O925f1 fragaria an
2	971	54.2	361	1	MTD_MBSCR	P93257 meembryant
3	963	54.8	337	1	MTD_PETCR	P42754 petroseilinu
4	960	54.6	357	1	MTD1_ARATH	Q02971 arabidopsis
5	959	54.6	359	1	MTD2_ARATH	Q02972 arabidopsis
6	946	53.8	365	1	MTD_APIGR	Q38707 apium grave
7	884	50.3	360	1	MTDH_ARATH	P42754 arabidopsis
8	870.5	49.5	359	1	MTD_MEDSA	Q82515 medicago sa
9	860.5	48.9	363	1	MTD3_STYHU	Q43138 stylosanthe
10	833.5	47.4	354	1	MTD1_STYHU	Q43137 stylosanthe
11	740	42.1	357	1	CAD2_PICAB	08205 picea abies
12		42.1	357	1	CAD7_PICAB	Q08350 picea abies
13	739	42.0	357	1	CADH_PINTA	P41637 pinus taeda
14	723	41.1	357	1	CADH_DINRA	Q40976 pinus radi
15	691.5	39.3	360	1	CAD1_ARACO	P42405 aralia cord
16	686	39.0	357	1	CAD9_TOBAC	P30360 nicotiana t
17	681	38.7	357	1	CAD4_TOBAC	P31359 nicotiana t
18	666	37.9	357	1	CADH_POPDE	P31657 populus del
19	656	37.3	361	1	CADH_LOLER	022380 loliu pere
20		37.2	356	1	CAD2_BUOGU	P31653 eucalyptus
21	653	37.1	356	1	CADH_BUOGU	064965 eucalyptus
22	646	36.7	357	1	CAD2_ARATH	049482 arabidopsis
23		36.5	358	1	CADH_MEDSA	P31656 medicago sa
24	629	35.8	354	1	CADI_EUCGU	042726 eucalyptus
25	628	35.7	367	1	CADH_RAIZE	Q44562 zea mays (m
26	621.5	35.4	355	1	CADI_BUCBO	P07046 eucalyptus
27	621	35.3	365	1	CADI_ARATH	P48523 arabidopsis
28	611	34.8	365	1	CADH_BACOF	082056 sacharum o
29	599.5	34.1	346	1	ADH_FYCTU	P31935 mycobacteri
30	555.5	31.6	349	1	YAHK_ECOLI	P75691 escherichia
31	379	21.6	360	1	YMG7 YEAST	Q04894 saccharomy
32	365.5	20.8	339	1	YJGB_BACLI	P42250 escherichia
33		20.3	339	1	YJGB_BACST	P42258 bacillus st

Query Match	62.1%; Score 1092; DB 1; Length 359;	Best Local Similarity	65.7%; Pred. No. 2 7e-82;	Matches	224; Conservative 27; Mismatches 58; Indels 32; Gaps 7;
Qy	1 ETGATDVRKFVLYCGVCHSDIHMKAQNDGTSIVTPIVPGHELYGVTEYVGCKVVKFKF-SWR 59	Db	32 ETGKDKDVTKFVLYCGVCHSDIHMVNRNEWGTSIVTPIVPGHBRIVETVEYVSKVQKVKVADR 91	Qy	60 QGRMTHLGRPTCENCINHLENCPNLIQTYGSKYDGTMTCGGSINMMATDEHETVRI 119
Db	92 VGCVICVGSCLR-SCENCNTHLENCPKQIILYGRAXYDGTSIVTPIVPGHBRIVETVEYVSKVQKVKVADR 150	Qy	120 PDNLPLDGAAPLLCAGITTSPPYYGLDKPGEALGVSVENPRFRSPRSPPLNPLGJNGSRLQ 179	Db	151 PDNLPLDGAAPLLCAGITTSPPYYGLDKPGEALGVSVENPRFRSPRSPPLNPLGJNGSRLQ 194
Qy	180 SLVPPLIKEGG---SYGTSR---ALMH---SLRITDODMEAAMSTMDGIIIDTVP 225	Db	195 HVVKFAKAMGKVKTIVTSPKKEERALKGADSPVRSRDQDMQAATGMDGIIIDTVS 254	Qy	226 AVRPLPLSILKETGKVTTGAVQPLDPLVPLIIGKVKMVAASSAIGE-KETOEIMDEA 285
Db	255 AQPPLPLGLNSRKGKLYMVGAPKPLPVPFLMGERKVKMVAASSGIGE-KETOEIMDEA 314	Qy	286 ARNITADLEVIPIDYLNAMERVKDVKYRFRFEDVNTL 326	Db	315 ARNITADLEVIPIDYLNAMERVLKADVYRFRFEDVNTL 355
Db	315 ARNITADLEVIPIDYLNAMERVLKADVYRFRFEDVNTL 355	RESULT 2	MTD_MESCR	STANDARD;	PRT; 361 AA.
AC	P32257;	AC	MTD_MESCR	STANDARD;	PRT; 337 AA.
DT	16-OCT-2001 (Rel. 40, Created)	DT	AC	P42754;	PRT; 337 AA.
DT	16-OCT-2001 (Rel. 40, Last sequence update)	DT	AC	P42754;	PRT; 337 AA.
DT	16-OCT-2002 (Rel. 42, Last annotation update)	DT	AC	P42754;	PRT; 337 AA.
DE	Probable mannitol dehydrogenase (EC 1.1.1.255)	DE	AC	P42754;	PRT; 337 AA.
DE	(NAD-dependent mannitol dehydrogenase).	DE	AC	P42754;	PRT; 337 AA.
GN	EL13.	GN	AC	P42754;	PRT; 337 AA.
OS	Mesembryanthemum crystallinum (Common ice plant).	OS	AC	P42754;	PRT; 337 AA.
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	OC	AC	P42754;	PRT; 337 AA.
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	OC	AC	P42754;	PRT; 337 AA.
OC	Caryophyllales; Aizoaceae; Mesembryanthemum.	OC	AC	P42754;	PRT; 337 AA.
OX	KCBI_TaxID=3544;	OX	AC	P42754;	PRT; 337 AA.
RN		RN	AC	P42754;	PRT; 337 AA.
RP		RP	AC	P42754;	PRT; 337 AA.
SEQUENCE FROM N.A.		SEQUENCE FROM N.A.	AC	P42754;	PRT; 337 AA.
RA	Michałowski C.B., Bohnert H.J.	RA	AC	P42754;	PRT; 337 AA.
RA	Submitted (Nov-1996) to the EMBL/GenBank/DDBJ databases.	RA	AC	P42754;	PRT; 337 AA.
CC	-!- FUNCTION: Oxidizes mannitol to mannose. Provides the initial step by which translocated mannitol is committed to central metabolism and, by regulating mannitol pool size, is important in regulating salt tolerance at the cellular level. (By similarity).	CC	AC	P42754;	PRT; 337 AA.
CC	-!- CATALYTIC ACTIVITY: D-mannitol + NAD(+) = D-mannose + NADH.	CC	AC	P42754;	PRT; 337 AA.
CC	-!- COFACTOR: Belongs to the zinc-containing alcohol dehydrogenase family.	CC	AC	P42754;	PRT; 337 AA.
CC	-!- CAUTION: Was originally (Ref.1) thought to be a cinnamyl-alcohol dehydrogenase.	CC	AC	P42754;	PRT; 337 AA.
CC	-!- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	CC	AC	P42754;	PRT; 337 AA.
CC	-!- CAUTION: Was originally (Ref.1) thought to be a cinnamyl-alcohol dehydrogenase.	CC	AC	P42754;	PRT; 337 AA.
CC	-!- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	CC	AC	P42754;	PRT; 337 AA.
DR	EMBL: U77071; AAB38503.1; -.	DR	AC	P42754;	PRT; 337 AA.
DR	DR	DR	AC	P42754;	PRT; 337 AA.
DR	InterPro: IPR002328; ADH zinc.	DR	AC	P42754;	PRT; 337 AA.
DR	InterPro: IPR002085; Adh_zinc family.	DR	AC	P42754;	PRT; 337 AA.
DR	Pfam: PF0107; Adh_zinc.	DR	AC	P42754;	PRT; 337 AA.
DR	PROSITE: PS00059; ADH_ZINC_1.	DR	AC	P42754;	PRT; 337 AA.

FT	METAL	68	68	ZINC 1	{CATALYTIC} (BY SIMILARITY).
FT	METAL	99	99	ZINC 2	(BY SIMILARITY).
FT	METAL	102	102	ZINC 2	(BY SIMILARITY).
FT	METAL	105	105	ZINC 2	(BY SIMILARITY).
FT	METAL	113	113	ZINC 2	(BY SIMILARITY).
FT	METAL	162	162	ZINC 1	{CATALYTIC} (BY SIMILARITY).
SQ	SEQUENCE	359	AA:	38942 MW:	B691F9B8AD4842AS CRC64;
Query Match		54.6%	Score 959;	DB 1;	Length 359;
Best Local Similarity		55.1%	Pred. No. 2.3e-71;		
Matches 193;	Conservative	39;	Mismatches	68;	Indels 50; G
Qy	1	ETGATDVRKFVLYCGVCHSDTEMAKNDWGTSTYPIVGHELVYVTEVGCKYKVKER			
Db	30	ETGEKDVRKFVLCIGIHCSDLEAVKQNEWGMSTYPLVPGHEIVGVWTBGAKYTKER			
Qy	61	-----GRCKLHGRAPTCNCIHLENKCNLQITYGSKYDGTMTYGGCSNNM-----			
Db	90	VGVGCLYSS-----GSDCSCTEGMENYCEKSIQTYGFPTDNTITYGGYSDM-----			
Qy	114	HITVRPINLPDGAAPLCAAGITTYSPWRYGLDKPFGMHLGVIEWRFRSRPPLN-----			
Db	143	GFVIRPDPNLPDAAAPLCAAGITVSKPDRKFGSLDGFGMHLGV-----			
Qy	174	WGSRLQLSIVPPPLIKEGGSYGTSPALMHS-----LIRTDODOMER-----			
Db	189	GG-----LGHGVYKEAKMGTKVTVISTSEKCRDEA1NRLGADAFVLSRDPKQIN-----			
Qy	217	MDGIIIDIVPAVRLPLEPLISLKLINGVTVGLAVQPLDLPVPLITGRMVAGSAT-----			
Db	244	MDGIIIDIVSATHSLPLIGLKLHKGKLNWGAPEKPLELPMPMLIETKRVYGMSSM-----			
Qy	277	ETOEMDFAAEHNTADIEVITDYLNTAMBERVVKDVRFRFVIDVENTL-----			326
Db	304	ETOEMDWAHGKNTADIELADISADYNTAMERKADYRVFVIDVANTL-----			353

RESULT 6					
MTD_APIGR	ID	MTD_APIGR	STANDARD;	PRT;	365 AA.
AC	Q38707	Q82461;			
DT	16-Oct-2001	(Rel. 40, Created)			
DT	16-Oct-2001	(Rel. 40, Last sequence update)			
DT	10-Oct-2003	(Rel. 42, Last annotation update)			
DE	Mannitol dehydrogenase	(EC 1.1.1.255)	(NAD-dependent mannitol dehydrogenase)		
MTD					
GN					
OS	Apium graveolens (Celery).				
OC	Pulvirota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; asterids; core eudicots; asterids; Apiales; Apiales; Apicidae; Apicidae; aploid superclade; Apium clade; Apium.				
OC	NCBI_TaxID=4045;				
OC					
OC					
OX					
RN	[1] -				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=9535325; PubMed=7638158;				
RA	Williamson J.D., Stoop J.M.H., Massei M.O., Conkling M.A., Pharr				
RT	"Sequence analysis of a mannitol dehydrogenase cDNA from plants reveals a function for the pathogenesis-related protein EL13.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 92:7148-7152(1995).				
RL	[2]				
RN	SEQUENCE FROM N.A.				
RA	Williamson J.D., Guo W.-W., Pharr D.M.;				
RT	"Cloning and characterization of a genomic clone encoding mannitol dehydrogenase from celery (Apium graveolens).";				
RT	[3] Plant Gene Register PG98-137.				
RN	[4]				
RP	CHARACTERIZATION				
RA	Stoop J.M.H., Chilton W.S., Pharr D.M.;				
RT	"Substrate specificity of the NAD-dependent mannitol dehydrogenase from celery.";				
RT	[5] Phytochemistry 43:1145-1150(1996).				
RL	[6]				
CC	- FUNCTION: Oxidizes mannitol to mannose. Provides the initial				

CC by which translocated mannitol is committed to central metabolism
 CC and by regulating mannitol pool size, is important in regulating
 CC salt tolerance at the cellular level.

CC -!- CATALYTIC ACTIVITY: D-mannitol + NAD(+) = D-mannose + NADH.

CC -!- COPARTNER: Binds 2 zinc ions per subunit (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC

CC DR U24561; AACI5467.1; -.

CC EMBL: AF0467082; AACG61854.1; -.

CC DR InterPro: IPR002328; ADH_zinc.

CC DR InterPro: IPR002085; Adh_zn_Family.

CC DR InterPro: IPR000205; NAD_BS.

CC DR InterPro: IPR000051; SAM_bind.

CC DR Pfam: PF00107; Adh_zinc_N; 1.

CC DR PROSITE: PS00059; ADH_ZINC; 1.

CC KW Oxidoreductase; Zinc; Metal-binding; NAD.

CC FT METAL 50 50 ZINC 1 (CATALYTIC) (BY SIMILARITY).

CC FT METAL 72 72 ZINC 1 (CATALYTIC) (BY SIMILARITY).

CC FT METAL 103 103 ZINC 2 (BY SIMILARITY).

CC FT METAL 106 106 ZINC 2 (BY SIMILARITY).

CC FT METAL 109 109 ZINC 2 (BY SIMILARITY).

CC FT METAL 117 117 ZINC 2 (BY SIMILARITY).

CC FT METAL 166 166 ZINC 1 (CATALYTIC) (BY SIMILARITY).

CC FT CONFLICT 47 47 C -> S (IN REF - 2).

CC FT CONFLICT 120 120 T -> I (IN REF - 2).

Query Match		Score 946: DB 1: Length 365;	
Best Local Similarity		56.5%: Pred. No. 2.7e-70;	
Matches 192; Conservative		42; Mismatches 74; Index 32; Gaps 5;	
QY	2	TGATDVREKVLKYGCHSDIEMAKNDWGTSITYPIVPGHBLVHTYEVGCKVKKEK-SWRQ	60
Db	35	TGKDVRLKVLKPGCVCSDEHMTHNNWCFSTYPIVPGHBLVHTYEVGCKVKVKGDNV	94
QY	61	GRCLLKGRLRPTCENCIHLLENCPNLQTYGSKVKYDSTMWGGYNNMVTDEHFLYRIP	120
Db	95	GIGLVGSCR-SCESSCCNRESHCENTDITYSIVFDSYTMHGGYSOTMVADEHFLRWP	153
QY	121	DNLPLDGAAPLLCAGITYSPWRYYGLDPGMELGVWPPRFRSRPFLNLPGWGSQIS	180
Db	154	KNPLDLDGAAPLLCAGITYSPWRYYGLDPGTGXIGV	197
QY	181	LVPPLIKRGGSYGTSPALMHS-----LIRTDDOMEAAMSTMGDIIDTYP	226
Db	198	VAYRMAKFGAQVTVIDTSESKRKEAELKLGAPELNSDQEMKGARSLSGDIIDTYPV	257
QY	227	VRPLEPLSILKTYKGKVYGLAQVPLDLPVFLLIGRKMKVAGSAIGGMKETOEMLDFAA	286
Db	258	NHPLAPLPLDKERKGKLYWGAPEKPKFLPVSLLGKRLKGTTGKETOEMLDFAA	317
QY	287	EHNITADEVIPDYLATAMERVKDVRFRVIDVENTL	326
Db	318	KINITADVEVIEVDYNTAMERLUVKSDYRYRFIDANTM	357
RESULT 7			
WTDB_ARATH	STANDARD	PRT;	360 AA.
ID	PA2734;		
AC			
DT	01-NOV-1995	(Rel. 32, Created)	
DT	01-NOV-1995	(Rel. 32, Last sequence update)	
DT	10-OCT-2003	(Rel. 42, Last annotation update)	
DT			

Probable mannitol dehydrogenase (EC 1.1.1.255) (NAD⁺-dependent mannitol dehydrogenase).
 CAD1 OR At4G39320 OR T22FB.230.
Arabidopsis thaliana (Mouse-ear cress);
 Eukarya; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; Spermatophytina; Magnoliophytina; eudicots; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis; Arabidopsis. NCBI_TaxID=3702;

[1] SEQUENCE FROM N.A.
 STRAIN=Cv. Columbia;
 STRAIN=Cv. IC
 MEDLINE=95357425; PubMed=7630954;
 Somers D.A., Nourse J.P., Manners J.M., Abrahams S.L., Watson J.M.; "A gene encoding a cinnamyl alcohol dehydrogenase homolog in *Arabidopsis thaliana*," *Plant Physiol.* 108:1309-1310 (1995).

[2] SEQUENCE FROM N.A.
 STRAIN=Cv. Columbia;
 MEDLINE=9803488; PubMed=10617198;
 Mayer K.F.X., Schueler C., Wambutt R., Murphy G., Volckaert G., Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N., Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M., Mueller M., Weichelhartgarter M., de Simonne V., Obermaier B., Machie R., Mueller M., Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T., Reichert B., Porteille D., Perez-Alonso M., Boultry M., Bancroft I., Oros P., Hoheisel J., Zimmermann W., Wedder H., Ridley P., Langham S.-A., McCullagh B., Bilham L., Robben J., "Vandenbussche F., van der Schueren J., Gromyronen B., Chuang Y.-J., Vandenbussche F., Braeck M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E., Weitzengger T., Botha G., Ramsperger U., Hilbert H., Braun M., Holzner B., Brandt A., Peters S., van Staveren M., Dirkse W., Moeljiman P., Klein Lankhorst R., Lyne M., Hauf J.J., Koetter P., Berneiser S., Hempt S., Feldpausch M., Lambeth S., Van den Daele H., De Keyser A., Buyschaert J., Gielen J., Villarreal K., De Clercq R., Van Montagu M., Rogers J., Cronicq A., Quail M.A., Bray-Alflen R., Clark L., Doggett J., Hall S., Kay M., Leonard N., McLAY K., Mayes R., Petetti A., Rajandream M.A., Lyne M., Benes V., Rechmann S., Borkova D., Bloedeker H., Schlarb M., Grimm M., Loehnert T.-H., Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Fartmann B., Grandjean K., Dauner D., Herzl A., Neumann S., Aszariou A., Vitale D., Ligouri R., Piravandi E., Massenet O., Quigley F., Clabaugh G., Muellein A., Felber R., Schnell S., Hiller R., Schmidt W., Lecharry A., Aubourg S., Scheffler F., Cooke R., Berger C., Monfort A., Casanbernita R., Gibbons T., Weber N., Varnelhol M., Bargues M., Terol J., Torres A., Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T., Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielek C., Friesman D., Haase D., Lemcke K., Mewes H.-W., Stocker S., Zaccaria P., Bevan R., Wilson R.K., de la Bastide M., Habermann K., Parmentier L., Dethia N., Gnoj L., Schutz K., Huang E., Spiegel L., Sekhon M., Murray J., Sheet P., Cordes M., Abu-Thraideh J., Stoenking T., Kalicki J., Graves T., Harmon G., Latreille P., Courtney L., Cloud J., Abbott A., Edwards J., Edwards K., Johnson D., Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K., Kemp K., Kramer J., Fulton L., Mardis R., Dante M., Pepin K., Hillier L.W., Nelson J., Spiech J., Ryan E., Andrews S., Geisel C., Layman D., Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C., Antoniou B., Zidiani M., Strong C., Sun H., Lanar B., Yordan C., Ma P., Zheng J., Preston R., Yil D., Hoffman M., Matero A., Shah R., Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S., Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A., Chen B., Narra M.A., Martiassen R., McCombie W.R.; "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana," *Nature* 402:769-777 (1999).

[3] FUNCTION: Oxidizes mannitol to mannose. Provides the initial step by which translocated mannitol is committed to central metabolism and, by regulating mannitol pool size, is important in regulating salt tolerance at the cellular level (By similarity).

[4] CATALYTIC ACTIVITY: D-mannitol + NAD(+) = D-mannose + NADH.

[5] COFACTOR: Binds 2 zinc ions per subunit (By similarity).

[6] SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenases.

198 HMAVKPAKTHEGLKITVISTSPPTKKEEAIRNLGADSEFLVSRDPQMEAPKETLDGILIDTV 257

225 PAVRPLEPLISLKLTKNGKVVTVGIAVOPFLIDLPVFPFLIGRMVAGSAIGGHKETOEMIDF 284
 Qy 258 SADHS1VPLIGLKLKSKHKLVLIGATEKPLELPPFLILGRKLVGGTLVGGKETQEMIDF 317
 Db 285 AAEHNITTADEEVIPIDYLNTAMERYVKKDVYRFRFIDVENTL 326
 Qy 318 SPKHNVPEEVVPMDVNIAQNLAKADVYRFRFIDVANTL 359
 Db

RESULT 10
 MTD1 STYHU STANDARD; PRT; 354 AA.
 AC Q43137;
 DT 15-DEC-1998 (Ref. 37, Created)
 DT 15-DEC-1998 (Ref. 37, Last sequence update)
 DT 10-OCT-2003 (Ref. 42, Last annotation update)
 DE Probable mannitol dehydrogenase 1 (EC 1.1.1.255) (NAD+-dependent
 DE mannitol dehydrogenase 1).
 GN NCBI_TaxID:35628;
 OS Stylosanthes humilis (Townsville stylo).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC Fabales; Fabaceae; Papilionoideae; Aescynomeneae;
 OC Stylosanthes.
 OX

RP SEQUENCE FROM N.A.
 RC STRAIN=Cv_Paterson; TISSUE=Stem; Manners J.M.; Curtis M.D.; Abrahams S.L.; Watson J.M.; Nourse J.P.; Manners J.M.; Curtis M.D.; Abrahams S.L.; Watson J.M.; RL Submitted (Nov-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Oxidizes mannitol to mannose. Provides the initial step by which translocated mannitol is committed to central metabolism and, by regulating mannitol pool size, is important in regulating salt tolerance at the cellular level (By similarity).
 CC -!- CATALYTIC ACTIVITY: D-mannitol + NAD(+) = D-mannose + NADH.
 CC -!- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
 CC -!- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase family.
 CC -!- CAUTION: Was originally (Ref.1) thought to be a cinnamyl-alcohol dehydrogenase.
 CC -!- CAUTION: This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -!- EMBL: L36823; AAA74982; -.
 DR InterPro; IPR002328; ADH_zinc.
 DR InterPro; IPR002085; Adh_zinc family.
 DR Pfam; PF00107; ADH_zinc_N; -.
 DR PROSITE; PS00059; ADH_ZINC; -.
 KW Oxidoreductase; Zinc; Metal-binding; NAD; Multigene family.
 FT METAL 43 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 65 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 96 ZINC 2 (BY SIMILARITY).
 FT METAL 99 ZINC 2 (BY SIMILARITY).
 FT METAL 102 ZINC 2 (BY SIMILARITY).
 FT METAL 110 ZINC 2 (BY SIMILARITY).
 FT METAL 158 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 SQ SEQUENCE 354 AA; 38130 MW; 662B93FD846034 CRC64;

Query Match 4 ADIVDREPKVLYCGVCESDIHAMKNDWTSITVIVPGHBLVGVVTEGCKVRFKSM-ROG-
 Best Local Similarity 49.0%; Fred. No. 4.5e-61;
 Matches 170; Conservative 47; Mismatches 79; Indels 51; Gaps 7;

Qy 30 ADDVTLKILMCGVCHSDLHTVQDNGFTTIVVFPGRHIAIGLTVTRGSNTVFKBGRDVGV 89
 Db

RESULT 11
 CAD2_PICAB ID CAD2_PICAB STANDARD; PRT; 357 AA.
 AC 082035; DT 16-OCT-2001 (Ref. 40, Created); DT 16-OCT-2003 (Ref. 42, Last sequence update); DE Cinnamyl-alcohol dehydrogenase 2 (EC 1.1.1.195) (CAD 2).
 OS Picea abies (Norway spruce) (Picea excelsa).
 OC Bokarota; Viridiplantae; Strptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
 RN [1] RR SEQUENCE FROM N.A.
 RA Schubert R.; Sperisen C.; Mueller-Starch G.; La Scala S.; Ernst D.; RA Sandermann H. Jr.; Haeger K.-E.; RR Karst: genomic sequences in *Picea abies* (L.) and phylogenetic relationships.";
 RL [1] RR FUNCTION: This protein catalyzes the final step in a branch of phenylpropanoid synthesis specific for production of lignin monomers. It acts on cinnamyl-, sinapyl-, 4-coumaryl- and
 CC -!- CATALYTIC ACTIVITY: Cinnamyl alcohol + NADP(+) = cinnamaldehyde + NADPH.
 CC -!- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
 CC -!- PATHWAY: Lignin biosynthesis
 CC -!- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase family.
 CC -!- CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -!- EMBL: AJ001924; CAA05095; 1; DR InterPro; IPR002328; Adh_zinc.
 DR InterPro; IPR002085; Adh_zinc family.
 DR Pfam; PF00107; ADH_zinc_N; -.
 DR PROSITE; PS00059; ADH_ZINC; -.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis; Multi-gene family.
 FT METAL 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 ZINC 2 (BY SIMILARITY).

FT METAL 114 124 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 SQ SEQUENCE 357 AA; 38834 MW; 5680133B933097FC CRC64;

Query Match Score 740; DB 1; Length 357;
 Best Local Similarity 42.1%; Pred. No. 2, 1e-53;
 Matches 147; Conservative 61; Mismatches 86; Indels 50; Gaps 5;

Qy 3 GATDVREKVLKYGCVCHSDIHMAKNDWGTSTYVTPGHELVGVYTEVSGKVKKKFWSRQ-- 60
 Db 33 GPDVTVRVIYICCHEDLWQHNEMGNSNTYMPGVHEVGVYTVTGEVKEKYGHEHVG 92
 Qy 61 ----GRCWLHGRRLPCTNCNTHLENKCPNLIQTCISKYDGTMTYGGSNMNTDDEHF 115
 Db 93 VCGIVGSC-----RCSNNGSM2YCSKPIWYDNPVHPTQGPGASMMVYDQMF 145
 Qy 116 TRIPTDNPFLDGAAPLICAGTTTSPRYRGLDKPENHGLGVWPRRFRSRSPPLNLPFLWG 175
 Db 146 VTRIPENPLQEAAPLICAGTTVSPKHFQTEPKGKCGI-----LGLGCVGH 194
 Qy 176 SRLQSLVPPPLKEGGSYGTSPALMHS-----LILRTDQDMEAAMSTMD 218
 Db 195 MG-----VJIAKAGLHLRTVISSSDRKREBALEYVGDAYLVSKDKEQAEASLD 246
 Qy 219 GLIDTPAVRPIEPLISLKLTKNGKVTYGLAQVPLDLPVPLIGKRMWASSAIGGMKET 278
 Db 247 YIMDTDVAHPPEPLKPLKNGKVLGIVYDPLHVTPLJLRLRSPISSEFIGSMEET 306
 Qy 279 QEMIDFAEHNTTADIEVPIPDYLNTAMERVVKDVRFRFVDD 322
 Db 307 QETLDFCAEKVSSMIEVVGGLDINTAMERLVNDYRYFVDD 350

RESULT 12
 CAD7_PICAB STANDARD; PRT; 357 AA.
 AC C08350; 31-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 19-OCT-2003 (Rel. 42, Last annotation update)
 DE Cinnamyl alcohol dehydrogenase 7/8 (EC 1.1.1.195) (CAD 7/8).
 GN CAD7 AND CAD8.
 OS Picea abies (Norway spruce) (Picea excelsa).
 OC Buxaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferales; Coniferales; Pinaceae; Picea.
 OC NCBI_TaxID:3329;
 RN [1] SEQUENCE FROM N.A.
 RP MEDLINE=90033286; PubMed=8219046;
 RA Galliano H., Cabane M., Eckerstorck C., Lottspeich F.,
 RA Sandermann H. Jr., Ernst D.;
 RT "Molecular cloning, sequence analysis and elicitor/ozone-induced
 accumulation of cinnamyl alcohol dehydrogenase from Norway spruce
 (Picea abies L.)";
 RL Plant Mol. Biol. 23:145-156(1993).
 RN [2] SEQUENCE FROM N.A.
 RA Schubert R., Sperisen C., Mueller-Starch G., La Scala S., Ernst D.,
 RA Sandermann H. Jr., Haeger K.-P.;
 RT "the cinnamyl alcohol dehydrogenase gene family in *Picea abies* (L.)
 and phylogenetic relationships. Southern hybridization, genetic analysis
 and phylogenetic relationships.";
 RL Trees 12:453-463 (1998).
 CC Pinerylpropanoid synthesis specific for production of lignin
 monomers. It acts on coniferyl-, sinapyl-, 4-coumaryl- and
 cinnamyl-alcohol.
 CC CATALYTIC ACTIVITY: Cinnamyl alcohol + NADP(+) = cinnamaldehyde +
 NADPH.
 CC COFACTOR: Binds 2 zinc ions per subunit (By similarity).
 CC PATHWAY: Lignin biosynthesis.
 CC SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase
 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to licenses@isb-sib.ch.)

CC -----
 DR EMBL; X72675; CAA51226; 1; -.
 DR EMBL; AJ001925; CAA05096; 1; -.
 DR EMBL; AJ001926; CAA05097; 1; -.
 DR PIR; S39509; S39509.
 DR InterPro; IPR002228; Adh_zinc.
 DR Pfam; PF00107; Adh_zinc_N_I.
 DR OXIDOREDUCTASE; Zinc; Metal-binding; NADP; Lignin biosynthesis;
 KW Multigene Family.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 SQ SEQUENCE 357 AA; 38777 MW; BCDPC41B7CA92 CRC84;
 Qy 3 GATDVRPKVLYCCVCHSDIHMAKNDWGTSTYVTPGHELVGVYTEVGCKVKKEPKSWRQ-- 60
 DR 33 GPEDVIRVYICGHSIDLVQENEMGSMEYQCSKRIWNTYDYNHDTPTQGFASSMVKDQF 115
 Qy 93 VGCIVGSC-----RCSNNGSM2YCSKPIWYDNPVHPTQGHELVGVYTEVGCKVKKEPKSWRQ-- 145
 DR 116 IVRIPDNPLDQPLGAPLICAGTTTSPRYGGLDKPQGHLGVEMPRRFRSRPPLNLPGLNG 175
 DR 146 VYRIPENPLEQDAPLICAGTTTSPRYGGLDKPQGHLGVEMPRRFRSRPPLNLPGLNG 194
 Qy 176 SRLQSLVPPPLIKEGGSYGTSPALMHS-----LILRTDQDMEAAMSTMD 218
 DR 195 MG-----VJIAKAGLHLRTVISSSDKKKEALEVIGADAYLVSKDKEQAEASLD 246
 Qy 219 GLIDTPAVRPIEPLISLKLTKNGKVTYGLAQVPLDLPVPLIGKRMWASSAIGGMKET 278
 DR 247 YMDTIPAHPLPEPLKPLKNGKVLGIVYDPLHVTPLJLRLRSPISSEFIGSMEET 306
 Qy 279 QEMIDFAEHNTTADIEVPIPDYLNTAMERVVKDVRFRFVDD 322
 DR 307 QETLDFCAEKVSSMIEVVGGLDINTAMERLVNDYRYFVDD 350

RESULT 13
 CADH_PINTA STANDARD; PRT; 357 AA.
 ID CADH_PINTA ID CADH_PINTA
 AC P41637; AC P41637;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE Cinnamyl-alcohol dehydrogenase (EC 1.1.1.195) (CAD).
 OS *Pinus taeda* (loblolly pine).
 OC Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC NCBI_TaxID:3352;
 RN [1] SEQUENCE FROM N.A.
 RC TISSUE_Xylem;
 RX MEDLINE=95327049; PubMed=7603432;

RA Mackay J.J., Liu W., Whetten R., Sederoff R., O'Malley D.;
 RT "Genetic analysis of cinnamyl alcohol dehydrogenase in loblolly pine: a single gene inheritance, molecular characterization and evolution.";
 RL Mol. Gen. Genet. 247:537-545 (1995).
 CC -: FUNCTION: This protein catalyzes the final step in a branch of phenylpropanoid synthesis specific for production of lignin monomers. It acts on coniferyl-, sinapyl-, 4-coumaryl- and cinnamyl-alcohols.
 CC -: CATALYTIC ACTIVITY: Cinnamyl alcohol + NADP(+) = cinnamaldehyde + NADPH.
 CC -: COFACTOR: Binds 2 zinc ions per subunit (By similarity).
 CC -: PATHWAY: Lignin biosynthesis.
 CC -: SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z37991; CBA86072; 1;
 DR EMBL: 237992; CBA86073; 1; -.
 DR PIR: S49443; S49444;
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38847 MW; 1B6051CA1CB57C19 CRC64;
 Query Match 42.0%; Score 739; DB 1; Length 357;
 Best Local Similarity 43.5%; Fred. No. 2.6e-53;
 Matches 147; Conservative 63; Mismatches 90; Indels 38; Gaps 5;
 QY 3 GATDVRFKVLQYGVCHSDIHAMKNDGTSPTPIVPHGELVYVTEVGCKYTKFKPSWQ--
 DB 33 GPDVITVIGICHSIDLQMRNENGMSHTPVSHVIVTEGSEYRKFKVGEHVG 92
 QY 61 -----GRCWLHGRLPRPTCENIHLHENYCPNLIQTYGSKYQDGTMYGGSSNNMVTDEHF 115
 DB 93 VGCIVGSC-----RSGCNGNQSMEQYCSRIVWNTNDHGTPQGTPASSWVYDQMF 145
 QY 116 IVRIPDNPLDGAAPLCAQGTTYSPTYGQLDKGPM-----HIGVEWPRFRS 164
 DB 146 VWRIPENPLPLEQAPLCAQTVPSMKHEAMTPESKKGCGLIGLGVGVLKAKAF-- 203
 QY 165 FPPNLPGLNGRSRQLSVPPLIKEGSSYGTSPALMELSLRPTQDOMEAMSTMDGJIDTV 224
 DB 204 -GLHVTVISSDQKKEAMEVTLGDAY-----LVSUDTEKRMBAESDLTMDT 252
 QY 225 PAVRPLEPLISLKTGKVYVQIPLDPLPFLIGRMVAGSAIGGMKETQEMIDE 284
 DB 253 FVAHFLPFLYPLALKTNGKLYVNLGVVPEPLHIVTPPLGRSLIASFIGMNEETQETLDF 312
 QY 285 AAERNHTADIVIPIDYLNTAMERVKKDVEFRFVVDV 322

Db 313 CAEKKVSMIEVVGLDINTAMERLENDVRYRFVVDV 350
 RESULT 14
 CADH_PINRA CADH_PINRA STANDARD; PRT; 357 AA.
 ID CADH_PINRA ID 081222;
 AC Q40976; 081222;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DB Cinnamyl-alcohol dehydrogenase (EC 1.1.1.195) (CAD).
 GN CAD.
 OS Pinus radiata (Monterey pine).
 OC Pinus radiata; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 CC NCBI_TAXID=3347;
 RN [1]
 RP SEQUENCE FROM N.A.
 CC TISSUE;Male cone;
 RA Wagner A.; Walden A.; Walter C.;
 RT "A cDNA encoding a cinnamyl alcohol dehydrogenase from Pinus radiata." (In) Plant Gene Register PGR96-097.
 RL [2]
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Moyle R.; Wagner A.; Walter C.;
 RT "Nucleotide sequence of a cinnamyl alcohol dehydrogenase (CAD) gene from Pinus radiata." (In) Plant Gene Register PGR96-118.
 RL [1]
 CC FUNCTION: This protein catalyzes the final step in a branch of lignin phenylpropanoid synthesis specific for production of lignin monomers. It acts on coniferyl-, sinapyl-, 4-coumaryl- and cinnamyl-alcohol.
 CC CATALYTIC ACTIVITY: Cinnamyl alcohol + NADP(+) = cinnamaldehyde + NADPH.
 CC COFACTOR: Binds 2 zinc ions per subunit (By similarity).
 CC PATHWAY: Lignin biosynthesis
 CC SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase family.
 CC
 DR EMBL: Z37991; CBA86072; 1;
 DR PIR: S49443; S49444;
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38847 MW; 1B6051CA1CB57C19 CRC64;
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AAB3874; 1;-
 DR EMBL: AP000491; AAC31166; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR EMBL: AP000491; AAC31166; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).
 FT METAL 114 114 ZINC 2 (BY SIMILARITY).
 FT METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT VARIANT 195 195 L > M.
 FT VARIANT 287 287 P > L.
 FT VARIANT 302 302 G > S.
 FT VARIANT 335 335 E > K.
 FT VARIANT 352 352 G > A.
 FT VARIANT 354 354 E > K.
 SQ SEQUENCE 357 AA; 38924 MW; CFB18429392D42D CRC64;
 DR EMBL: U62394; 1;-
 DR InterPro: IPR002328; ADH_zinc.
 DR InterPro: IPR002085; Adh_zn_family.
 DR Pfam: PF00107; ADH_zinc_N; 1.
 DR PROSITE: PS00059; ADH_ZINC; 1.
 KW Oxidoreductase; Zinc; Metal-binding; NADP; Lignin biosynthesis.
 FT METAL 47 47 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).
 FT METAL 100 100 ZINC 2 (BY SIMILARITY).
 FT METAL 103 103 ZINC 2 (BY SIMILARITY).
 FT METAL 106 106 ZINC 2 (BY SIMILARITY).

Qy	61	-----GRCWHLGRRLPTCENCICHHLENQYCNLIIQTYGSKYDGMWMTGGYSNNKVTDEHF	115	-----	SQ	SEQUENCE	360	AA;	39129	MW;	28FD980E08C4096	CRC64;
Db	93	VGGIVGSC-----RSCTPDLQAPLCAAGTTSPPWRYGLDKPGMHGIVEMPERFSSPPLNLPGLWG	145	-----	Qy	39.1%; Score: 691.5; DB 1; Length: 360; Best Local Similarity 44.0%; Pred. No. 2.1e-49; Matches 52; Mismatches 100; Indels 39; Gaps 9;						
Qy	116	TVRIPDNPLPDLQAPLCAAGTTSPPWRYGLDKPGMHGIVEMPERFSSPPLNLPGLWG	175	-----	Qy	1 ETGATGTVRFKVLYCGVCHSDIHMAKNDKGTRSTYPIVPGHELVGVYKVKPFKSWRQ 60						
Db	146	VVRTPENPLPQAAAPLCAAGTTSPPMKGKAMTEPGKKG1-----LGIGVGH 194	-----	-----	Db	1 ETGATGTVRFKVLYCGVCHSDIHMAKNDKGTRSTYPIVPGHELVGVYKVKPFKSWRQ 60						
Qy	176	SRLQSLVPPPLIKEGGSGTSPALMHS-----LIRTDQDOMEAMSTND 218	-----	-----	Db	31 ETGCPENFPIKIIYCGCHTDQIANDLGASNYPRPGHEVGEVGSIVTKF---V 87						
Db	195	MG-----VKAIAFGIIGTVTWSSSDKKEEAEMLVGDAYLVSQDKMMPAESLD 246	-----	-----	Qy	61 GRCWHLGRL---REPTENCTIHLHENYCPNLIQTYGSKYDGMWMTGGYSNNKVTDEHFIV 11.7						
Qy	219	GIDITVPAVPPLEPLISLJKTNGKWTGIAVQPLDLPVPLIGRKAWSAIGKRET 278	-----	-----	Db	88 GDCVGGDTIVGCKKPRCPKADVEOCNKCIWSFDDVYTGKPTOGFESGRMVYDOKFVW 14.7						
Db	247	YIMITVPAHPLPEPLAALVPLAHEVPLIGRPSAIGKRET 306	-----	-----	Qy	118 RPPDNPLQDPLAPLICAGTTSPPRYYGLDK-PGMLHGVWPKRPRSRSPRPNLPGWGS 176						
Qy	279	QEMIDFAAEHHNTTADIEVIFDYLNTAMERVKKDVRFRFVIDV 322	-----	-----	Db	148 K1PDRGAPEQAPLICAGTVVSPTHFGLKEISIRGG1-----LGLGEGVGHM 19.6						
Db	307	QETLDFCAEKXKSSMIEVVGLDYINTAMELFLKNDVRYFVVDV 350	-----	-----	Qy	177 RLQSELVPLTKEGGYGT-----SLLRRTDQDQMEAMAMSTNDGIIID 222						
RESJLT 15					Db	197 GVK---LAKAMGHTWVSSSDKKEBDIDLGMDTIVSSDATMQZADSIDYIID 251						
CAD1_ARACO					Qy	223 TYPAVRPLEPLISLKTNSKRYVTVGLAVQPLDLPVFLI-IGRKAHVGASAGGMKETQEM 281						
ID CAD1_ARACO					Db	252 TYPVPHPLEPLVSLKLQSKLILMGVINTPLQF-SPSMWMLGKATGSPFGSMKETEEM 310						
AC P42435;					Qy	282 IDPNAHNNTADIBTPIKXLTAMEVVKDVRFRFVIDV 322						
DT 01-NOV-1995					Db	311 LDFCNEKGITSTIEVVKDYNTAFTERLKDNDVYRFVVDV 351						
DT 01-NOV-1995		(Rel. 3.2, Created)										
DT 10-OCT-2003		(Rel. 3.2, Last sequence update)										
DE CADD1		Cinnamyl-alcohol dehydrogenase 1 (EC 1.1.1.195) (CAD).										
GN OS		Aralia cordata (Udo) (Cordate spikenard).										
OC SPERMATOPHYTA		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;										
OC CAMPAENALYDIA		Magnoliophyta; eudicots; core eudicots; asterids;										
NCBI_TAXID=29746;		Campanulales; Apiales; Arales; Arales; Arales; Arales.										
RN												
RP		SEQUENCE FROM N.A.										
RA		RA										
RT		RT										
RT		"Cinnamyl alcohol dehydrogenase from <i>Aralia cordata</i> : cloning of the cDNA and expression of the gene in lignified tissues."										
RL		Plant Cell Physiol. 34:659-665 (1993).										
CC		-!- FUNCTION: This protein catalyzes the final step in a branch of										
CC		phenylpropanoid synthesis specific for production of lignin										
CC		monomers. It acts on coniferyl-, sinapyl-, 4-coumaryl- and										
CC		cinnamyl-alcohol.										
CC		-!- CATALYTIC ACTIVITY: Cinnamyl alcohol + NADP(+) = cinnamaldehyde + NADPH.										
CC		-!- COFACTOR: Binds 2 zinc ions per subunit (By similarity).										
CC		-!- PATHWAY: Lignin biosynthesis										
CC		-!- SIMILARITY: Belongs to the zinc-containing alcohol dehydrogenase family.										
CC												
DR		DR										
DR		InterPro; IPR02328; Adh_zinc.										
DR		InterPro; IPR002085; Adh_zn family.										
DR		PROSITE; PS00059; ADH_ZINC_1.										
KW		Oxidoreductase; Zinc; metal-binding; NADP; Lignin biosynthesis.										
FT		ZINC 1 (CATALYTIC) (BY SIMILARITY).										
FT		METAL 69 69 ZINC 1 (CATALYTIC) (BY SIMILARITY).										
FT		METAL 100 100 ZINC 2 (BY SIMILARITY).										
FT		METAL 106 106 ZINC 2 (BY SIMILARITY).										
FT		METAL 114 114 ZINC 2 (BY SIMILARITY).										
FT		METAL 163 163 ZINC 1 (CATALYTIC) (BY SIMILARITY).										

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: D13991; BA030991; -.

DR InterPro; IPR02328; Adh_zinc.

DR PROSITE; PS00059; ADH_ZINC_1.

PFAM: PF00107; ADH_zinc_N; 1.

ZINC 1 (CATALYTIC) (BY SIMILARITY).

ZINC 1 (CATALYTIC) (BY SIMILARITY).

ZINC 2 (BY SIMILARITY).

ZINC 2 (BY SIMILARITY).

ZINC 2 (BY SIMILARITY).

ZINC 1 (CATALYTIC) (BY SIMILARITY).

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2004, 13:33:14 ; Search time 31.8458 Seconds
Perfect score: 1758
Sequence: 1 BTGATTGVRKVLYCGVCHSD.....ERVVKDVRFREVIDVENTL 326

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : SPTREML25:*

1: sp_archae:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rabbit:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

17 679 38.6 357 10 Q9M632 populus tre
18 675 38.4 357 10 Q9FSC7 populus tri
19 666 37.9 407 10 Q8s412 loliun pere
20 662 37.7 361 10 Q94781 festuca aru
21 660 37.5 361 10 Q94783 festuca aru
22 660 37.5 361 10 Q94782 festuca aru
23 654 37.2 361 10 Q94780 festuca aru
24 648 36.9 356 10 Q9FUB8 eucalyptus
25 641 36.5 335 10 Q8W420 medicago sa
Q9ae96 mycobacteri
Q88k5 pseudomonas
Q884b3 pseudomonas
Q88krc3 myxococcus
Q9ulfo leishmania
Q92md4 rhizobium m
Q8xqn4 ralstonia s
Q88k65 arabiobacteri
Q82144 streptomyce
Q9cbq3 mycobacteri
Q04079 zinnia eleg
Q8h859 oryza sativ
Q8uf43 arabiobacteri
Q8l9u1 bacillus su
Q88241 pseudomonas
Qaess57 oceanobacil
Q9ca33 arabiobacteri
Q8pr22 xanthomonas
Q9ijj9 pseudomonas
Q8y1q7 ralstonia s

ALIGNMENTS

RESULT 1

Q9ATW1	ID	Q9ATW1	PRELIMINARY	PRT
AC	AC	Q9ATW1;		
DT	DT	01-JUN-2001 (TRIMBUREL, 17, Created)		
DT	DT	01-JUN-2001 (TRIMBUREL, 17, Last sequence update)		
DT	DT	01-OCT-2003 (TRIMBUREL, 17, Last annotation update)		
DB		Cinnamyl alcohol dehydrogenase		
OS		Fragaria ananassa (Strawberry)		
OC		Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; euroids I; Rosales; Rosaceae; Rosoideae; Fragaria.		
OC		NCBI_TaxID=3747;		
RN		RN		
RP		RP		
RC		SEQUENCE FROM N.A.		
RA		STRAIN=cv. Chandler; Blanco-Portales R.R., Caballero-Capulio J.J., Munoz-Blanco J.J.;		
RT		"Cloning, expression and immunolocalization pattern of a cinnamonyl alcohol dehydrogenase gene from strawberry (Fragaria x ananassa c.v. Chandler)." ;		
RL		Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.		
CC		EMBL: AF280509.1; -		
CC		-; COFACTOR: ZINC (BY SIMILARITY).		

RT "Cloning, expression and immunolocalization pattern of a cinnamonyl alcohol dehydrogenase gene from strawberry (Fragaria x ananassa c.v. Chandler)." ;

RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR DR GO; GO:0004024: F: alcohol dehydrogenase activity, zinc-dependent, IEA.

DR DR GO; GO:0005489: F: oxidoreductase activity; IEA.

DR DR GO; GO:0016491: F: zinc ion binding; IEA.

DR DR GO; GO:0008270: F: zinc ion binding; IEA.

DR DR GO; GO:0006118: P: electron transport; IEA.

DR DR InterPro: IPR02328: ADH_zinc.

DR DR InterPro: IPR02085: Adh_zn_family.

DR DR GO; GO:0005489: F: cytochrome c, zinc-dependent, IEA.

DR DR Pfam: PF00107; ADH_zinc_N_1.

DR DR PROSITE; PS00059; ADH_ZINC_1.

DR DR PROSITE; PS00190; CYTOCHROME_C_1.

KW Metal-binding; Oxidoreductase; Zinc.

SQ SEQUENCE 359 AA; 38907 MW; 5432F24RA3C64; CRC64;

Query Match 60.5%; Score 1064; DB 10; Length 359; Best Local Similarity 64.5%; Pred. No. 6.1e-89;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match	Length	DB	ID	Description
1	09ATW1	359	10	Q9ATW1		09atwl fragaria an
2	Q94GS9	362	10	Q94GS9		Q94959 populus tre
3	Q7XB2	360	10	Q7XB2		Q7xb2 camptotheaca
4	Q8H0U8	359	10	Q8H0U8		Q8h118 solanum tub
5	Q8L7U8	362	10	Q8L7U8		Q8l7u8 populus tre
6	Q94K02	360	10	Q94K02		Q94a02 arabidopsis
7	Q8LB84	360	10	Q8LB84		Q8l184 arabidopsis
8	Q85621	363	10	Q85621		Q85621 arabidopsis
9	Q8S411	370	10	Q8S411		Q8s11 loliun pere
10	Q9S110	375	10	Q9S110		Q9s110 arabidopsis
11	Q9S125	376	10	Q9S125		Q9s125 arabidopsis
12	Q7XB98	420	10	Q7XB98		Q7x98 oryza sativ
13	Q9M722	289	10	Q9M722		Q9m722 lycopersico
14	Q7XWU0	363	10	Q7XWU0		Q7xwu0 oryza sativ
15	Q8HB09	391	10	Q8HB09		Q8h109 oryza sativ
16	Q7XLDS	410	10	Q7XLDS		Q7xlds5 oryza sativ

Matches	220;	Conservative	27;	Mismatches	62;	Indels	32;	Gaps	7;
Qy	1	ETGATDYEKPLYCGVCHSDIMAKNDGCTTSPYIPIGHELYGVVTEVGCKKUFKEK-SWR	59						
Db	32	ETGKDYMVKYLGICHSIDHMVKYNEWGFTSPYIPIGHELYGVVTEVGCKKUFKEK-VGDR	91						
Qy	60	QGRCWHLGRLEPTCBN1HLLENYCPMLQYGSKTYDGMITYGGYSNNMFTDEHFTIVRI	119						
Db	92	VGVGCVVGCSR-SCENCTDHLENYCPQIQLTGGANTYDGTITGGCDIMVAHEHFFVRI	150						
Qy	120	PDNLPLDGAAPLCAAGTTSPWRYGLDKPQGMHLYGWEPRRSRPLNLPGLMGSRLQ	179						
Db	151	PDNLPLDGAAPLCAAGTTSPWRYGLDKPQGMHLYGWEPRRSRPLNLPGLMGSRLQ	179						
Qy	180	SLVPLIKEGG---SYGTSP---ALMH---SILRTDODOMEAMSTMDGTTDTPV	225						
Db	195	HVAVKFAKAMCAGVKVTVIYSTSKEEBAKHLGADSPFLVSRDQDHMORAIGTMDGTTV	254						
Qy	226	AVRPLPLESLIQLTKGKVTVGLAVQPLDPIPLIGRQVAGSATGGMKETQEMIDFA	285						
Db	255	AQHPLPLIGLKLXSHGKLYWGAPEKPLFPPLMGRQVAGSGIGMMETQEMIDFA	314						
RESULT 3									
	294G59	PRELIMINARY;							
AC	Q94G59;	PRELIMINARY;							
DT	01-DEC-2001	(TREMBLrel. 19, Created)							
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)							
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)							
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)							
RN	[1]								
SEQUENCE FROM N.A.									
RA	Gorman E.B., McKnight T.D.								
RT	"The origin of 10-hydroxygeraniol oxidoreductase."								
RL	Submitted (JUL-2003) (TREMBLrel. 25, Last sequence update)								
DR	EMBL; AY342335; AAQ20892.1; -.								
SQ	SEQUENCE 360 AA; 44BB7C85SCA49E5 CRC64;								
Query Match 55.2%; Score 570; DB 10; Length 360;									
Best Local Similarity 54.2%; Pred. No. 2.5e-80; Gaps 5;									
Matches 186; Conservative 57; Mismatches 62; Indels 38;									
Qy	2	TGATDVRFRKVLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Db	35	TGEFDVKFVPMCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Query Match 61 GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP 120									
Qy	61	GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	120						
Db	95	GVGCMVGSR-SCNDCDNN-ENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	153						
Qy	121	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Db	154	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Qy	181	LVPPLIKEGGSYGTSPALMHS-----LIRTDQMEAMSTMDDIJD	223						
Db	195	LGHVAKVPAKALGKVTVISTSLAKKEAERLGADSEFLYSDTQMAQAGTLLD	254						
Qy	224	VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID	283						
Db	255	VSATHEPLPLGLKSHGRLVLYGAPEKPLWPLMGRKIVGSKNNMVKYQEMID	314						
Query Match 224 VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID									
Best Local Similarity 57.2%; Pred. No. 1.3e-81; Gaps 6;									
Matches 198; Conservative 37; Mismatches 67; Indels 44;									
Qy	2	TGATDVRFRKVLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Db	35	TGEEBDYREKPLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Query Match 61 GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP 120									
Qy	61	GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	120						
Db	95	GVGCMVGSR-SCNDCDNN-ENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	153						
Qy	121	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Db	154	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Qy	181	LVPPLIKEGGSYGTSPALMHS-----LIRTDQMEAMSTMDDIJD	223						
Db	195	LGHVAKVPAKALGKVTVISTSLAKKEAERLGADSEFLYSDTQMAQAGTLLD	254						
Qy	224	VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID	283						
Db	255	VSATHEPLPLGLKSHGRLVLYGAPEKPLWPLMGRKIVGSKNNMVKYQEMID	314						
Query Match 224 VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID									
Best Local Similarity 57.2%; Pred. No. 1.3e-81; Gaps 6;									
Matches 198; Conservative 37; Mismatches 67; Indels 44;									
Qy	2	TGATDVRFRKVLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Db	35	TGEEBDYREKPLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Query Match 61 GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP 120									
Qy	61	GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	120						
Db	95	GVGCMVGSR-SCNDCDNN-ENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	153						
Qy	121	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Db	154	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Qy	181	LVPPLIKEGGSYGTSPALMHS-----LIRTDQMEAMSTMDDIJD	223						
Db	195	LGHVAKVPAKALGKVTVISTSLAKKEAERLGADSEFLYSDTQMAQAGTLLD	254						
Qy	224	VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID	283						
Db	255	VSATHEPLPLGLKSHGRLVLYGAPEKPLWPLMGRKIVGSKNNMVKYQEMID	314						
Query Match 224 VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID									
Best Local Similarity 57.2%; Pred. No. 1.3e-81; Gaps 6;									
Matches 198; Conservative 37; Mismatches 67; Indels 44;									
Qy	2	TGATDVRFRKVLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Db	35	TGEEBDYREKPLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Query Match 61 GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP 120									
Qy	61	GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	120						
Db	95	GVGCMVGSR-SCNDCDNN-ENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	153						
Qy	121	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Db	154	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Qy	181	LVPPLIKEGGSYGTSPALMHS-----LIRTDQMEAMSTMDDIJD	223						
Db	195	LGHVAKVPAKALGKVTVISTSLAKKEAERLGADSEFLYSDTQMAQAGTLLD	254						
Qy	224	VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID	283						
Db	255	VSATHEPLPLGLKSHGRLVLYGAPEKPLWPLMGRKIVGSKNNMVKYQEMID	314						
Query Match 224 VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID									
Best Local Similarity 57.2%; Pred. No. 1.3e-81; Gaps 6;									
Matches 198; Conservative 37; Mismatches 67; Indels 44;									
Qy	2	TGATDVRFRKVLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Db	35	TGEEBDYREKPLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Query Match 61 GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP 120									
Qy	61	GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	120						
Db	95	GVGCMVGSR-SCNDCDNN-ENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	153						
Qy	121	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Db	154	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Qy	181	LVPPLIKEGGSYGTSPALMHS-----LIRTDQMEAMSTMDDIJD	223						
Db	195	LGHVAKVPAKALGKVTVISTSLAKKEAERLGADSEFLYSDTQMAQAGTLLD	254						
Qy	224	VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID	283						
Db	255	VSATHEPLPLGLKSHGRLVLYGAPEKPLWPLMGRKIVGSKNNMVKYQEMID	314						
Query Match 224 VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID									
Best Local Similarity 57.2%; Pred. No. 1.3e-81; Gaps 6;									
Matches 198; Conservative 37; Mismatches 67; Indels 44;									
Qy	2	TGATDVRFRKVLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Db	35	TGEEBDYREKPLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Query Match 61 GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP 120									
Qy	61	GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	120						
Db	95	GVGCMVGSR-SCNDCDNN-ENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	153						
Qy	121	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Db	154	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Qy	181	LVPPLIKEGGSYGTSPALMHS-----LIRTDQMEAMSTMDDIJD	223						
Db	195	LGHVAKVPAKALGKVTVISTSLAKKEAERLGADSEFLYSDTQMAQAGTLLD	254						
Qy	224	VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID	283						
Db	255	VSATHEPLPLGLKSHGRLVLYGAPEKPLWPLMGRKIVGSKNNMVKYQEMID	314						
Query Match 224 VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID									
Best Local Similarity 57.2%; Pred. No. 1.3e-81; Gaps 6;									
Matches 198; Conservative 37; Mismatches 67; Indels 44;									
Qy	2	TGATDVRFRKVLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Db	35	TGEEBDYREKPLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Query Match 61 GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP 120									
Qy	61	GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	120						
Db	95	GVGCMVGSR-SCNDCDNN-ENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP	153						
Qy	121	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Db	154	DNTPLDGAAPLCAAGTTSPWRYGSDKGMLHGLCNEWPRRSRPLNPLGLWSRLQS	180						
Qy	181	LVPPLIKEGGSYGTSPALMHS-----LIRTDQMEAMSTMDDIJD	223						
Db	195	LGHVAKVPAKALGKVTVISTSLAKKEAERLGADSEFLYSDTQMAQAGTLLD	254						
Qy	224	VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID	283						
Db	255	VSATHEPLPLGLKSHGRLVLYGAPEKPLWPLMGRKIVGSKNNMVKYQEMID	314						
Query Match 224 VPAVREPLPLESLIQLTKGKVTVGLAVQPLDLPVPLIGRQVAGSATGGMKETQEMID									
Best Local Similarity 57.2%; Pred. No. 1.3e-81; Gaps 6;									
Matches 198; Conservative 37; Mismatches 67; Indels 44;									
Qy	2	TGATDVRFRKVLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Db	35	TGEEBDYREKPLYCGVCHSDIHMARNQWGTSTYPIVGHEDLVVTEVGKVRKFK-SWRQ	60						
Query Match 61 GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKYDGMITYGGYSNNMVTDEHFIVRIP 120									
Qy	61	GRCMLHGRLRPTCENCIIHLHENYCPNLIIZYGSKY							

AC	Q8H018;	01-MAR-2003	[TREMBLrel. 23, Created]	RN	RP	SEQUENCE FROM N.A.
DT		01-MAR-2003	(TREMBLrel. 23, last sequence update)	RA	RA	Israelsson M., Eriksson M.E., Hertzberg M., Aspeborg H., Nilsson P., Moritz T.;
DT		01-OCT-2003	(TREMBLrel. 25, last annotation update)	RT	RT	"Charges in gene transcription in the wood-forming tissue of transgenic hybrid trunks with increased secondary growth."
GN				RL	RL	Submitted (JGN-2002) to the EMBL/GenBank/DBJ databases.
DN				EMBL:	EMBL:	RT126444; AAM95578.1;
GN				DR	DR	GO; GO:0004024; Fructose dehydrogenase activity; zinc-dependent, IEA.
OS				DR	DR	GO; GO:0005489; Fr:electron transporter activity; IEA.
OS				DR	DR	GO; GO:0008270; Zinc ion binding; IEA.
OC				DR	DR	GO; GO:0006118; Zinc ion transport; IEA.
OC				DR	DR	InterPro; IPK02348; Adh_zinc.
OC				DR	DR	InterPro; IPK00205; Adh_zn_family.
OC				DR	DR	InterPro; IPK00345; CytC_heme_BS.
OC				DR	DR	PTM; PF00107; Adh_zinc_N_1.
OC				DR	DR	PROSITE; PS00059; Adh_zinc_C_1.
OC				DR	DR	PROSITE; PS00190; CYTOCHROMC_1.
OX				SO	SO	SEQUENCE 362 AA; 39527 MW; BA512D2CBFC97EC6 CRC64;
OX				Query Match	Query Match	Score 936; DB 10; Length 362;
OX				Best Local Similarity	Best Local Similarity	Score 936; DB 10; Length 362;
OX				Matches	Matches	Pred. No. 3.3e-77; Indels 36; Gaps 3;
OX				1	1	ETGATDVRFKVLGYCGVCSEDIAMRNIDGSTTYPVPGHRLGVVTEVGKVKPFKSWRQ 60
OX				2	2	TGATDVRFKVLGYCGVCSEDIAMRNIDGSTTYPVPGHRLGVVTEVGKVKPFKSWRQ 60
OX				3	3	DNGVEDVTKILYCGVCSEDIAMRNIDGSTTYPVPGHRLGVVTEVGKVKPFKSWRQ 60
OX				4	4	DNGVEDVTKILYCGVCSEDIAMRNIDGSTTYPVPGHRLGVVTEVGKVKPFKSWRQ 60
OX				5	5	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				6	6	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				7	7	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				8	8	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				9	9	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				10	10	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				11	11	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				12	12	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				13	13	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				14	14	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				15	15	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				16	16	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				17	17	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				18	18	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				19	19	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				20	20	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				21	21	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				22	22	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				23	23	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				24	24	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				25	25	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				26	26	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				27	27	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				28	28	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				29	29	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				30	30	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				31	31	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				32	32	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				33	33	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				34	34	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				35	35	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				36	36	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				37	37	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				38	38	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				39	39	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				40	40	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				41	41	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				42	42	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				43	43	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				44	44	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				45	45	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				46	46	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				47	47	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				48	48	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				49	49	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				50	50	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				51	51	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				52	52	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				53	53	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				54	54	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				55	55	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				56	56	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				57	57	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				58	58	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				59	59	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				60	60	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				61	61	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				62	62	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				63	63	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				64	64	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				65	65	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				66	66	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				67	67	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				68	68	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				69	69	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				70	70	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				71	71	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				72	72	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				73	73	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				74	74	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				75	75	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				76	76	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				77	77	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				78	78	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				79	79	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				80	80	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				81	81	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				82	82	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				83	83	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				84	84	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				85	85	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				86	86	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				87	87	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				88	88	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				89	89	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				90	90	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				91	91	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				92	92	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				93	93	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				94	94	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				95	95	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				96	96	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				97	97	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				98	98	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				99	99	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				100	100	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				101	101	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				102	102	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				103	103	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				104	104	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				105	105	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				106	106	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				107	107	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				108	108	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				109	109	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				110	110	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				111	111	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				112	112	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				113	113	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				114	114	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				115	115	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				116	116	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				117	117	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				118	118	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				119	119	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				120	120	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				121	121	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				122	122	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				123	123	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				124	124	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				125	125	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				126	126	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				127	127	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				128	128	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				129	129	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				130	130	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				131	131	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				132	132	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				133	133	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				134	134	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				135	135	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				136	136	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				137	137	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				138	138	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX				139	139	GRWLGHLGRURPCTENCIELENVCPNLQIOTYGSKYDGTMTYGGSYNNMTDEHFLVRIP 120
OX		</td				

RN	SEQUENCE FROM N.A.	RN	SEQUENCE FROM N.A.
RP	Lam B., Southwick A., Karlin-Newmann G., Nguyen M., Miranda M.,	RP	SEQUENCE FROM N.A.
RA	Palm C.J., Bowser L., Jones T., Bath J., Carninci P., Chen H.,	RA	Brower V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA	Cheuk R., Chung M.K., Hayashizaki Y., Ishii K., Kamiya A., Kawai J.,	RA	Feldmann K.;
RA	Kim C., Lin J., Liu S., Naruseka M., Pham P.K., Sakano H.,	RA	"Full-length cDNA from Arabidopsis thaliana,"
RA	Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,	RA	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RA	Becker J., Theologis A., Davis R.W.;	RA	EMBL; AY087363; F:alcohol dehydrogenase activity, zinc-dependent; IEA.
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.	DR	GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.
RN	SEQUENCE FROM N.A.	DR	GO; GO:00040489; F:electron transporter activity; IEA.
RP	Kim S.J., Kim M.R., Bedgar D.L., Moinuddin S.G.A., Cardenas C.L.,	DR	GO; GO:00040270; F:electron transport; IEA.
RA	Davin L.B., Lewis N.G.;	DR	InterPro; IPR020851; Adh_zn_family.
RA	Functional Reclassification of the Putative Cinnamyl Alcohol	DR	InterPro; IPR00345; Cyt_C_heme_BS.
RT	Dehydrogenase (CAD) Multigene Family in Arabidopsis ";	DR	InterPro; IPR0107; Adh_zinc_N_1.
RT	Dehydrogenase (CAD) Multigene Family in Arabidopsis ";	DR	PROSITE; PS00059; Adh_2_ZNC_1.
RL	Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.	DR	PROSITE; PS00190; CYTOCHROME_C_1.
DR	EMBL; AY30498; AAK4375..1..;	DR	SEQUENCE 360 AA; 06DD3B0581785759 CRC64;
DR	EMBL; AY302976; AAU47376..1..;	DR	Query Match 1 ETGATDVRPKLYCCCHSDHIMAKNDWGRSTYPTVPGHLYGVYTEVGCKYKKPKSWRQ 60
DR	GO; GO:00040469; AAU47376..1..;	DR	Best Local Similarity 48.9%; Pred. No. 1e-71; Indels 50; Gaps 5;
DR	PIR; S71179; S71179.	DR	Matches 171; Conservative 52; Mismatches 77; Indels 171; Gaps 5;
DR	PIR; T08581; T08581.	DR	34 DNGENDVTKILFGCVTDLHTKWDGTSYTFPPGHDVIGATKVGNVTRPKGDR 93
DR	GO; GO:0038270; F:zinc ion binding; IEA.	DR	61 -----GRCWLHGRRLRPTCENCIFHLENYCPLQITYGSKYDSTMVCGSYNNMVYDDE 113
DR	GO; GO:0006118; P:electron transport; IEA.	DR	94 VGVGVISGSC-----QSCESCPDQDLBNYCPNSCTYNSGDSDTRKNSYNSVVDQ 146
DR	InterPro; IPR002328; Adh_zinc.	DR	114 HETRIPDNFLDGAAPLCLAGITVSPWRYGLDQKPMAMHSPVRRFRSRPPLNPGL 173
DR	InterPro; IPR02085; Adh_zn_family.	DR	147 RFLVRLPENLPSDSCAAPLCLAGITVSPWRYGLDQKPMAMHSPVRRFRSRPPLNPGL 173
DR	InterPro; IPR00345; Cyt_C_heme_BS.	DR	174 WGSRLQSLVPLPLKCGSYGTSPALMHS-----LRLRQDQMEAMST 216
DR	PFAM; PF00107; Adh_zinc_N_1.	DR	193 GG-----LGHAVVIGKQASLKVYISSSSSTKAEAHINHISADSLTVTIDQKMKAKIGT 247
DR	PROSITE; PS00059; Adh_zinc_N_1.	DR	217 MDGIIIDTVPLRPLISLKLTKVWVYQIOPDLFWFPLIILGRKVAGSAIGGMK 276
DR	COXIDREDUCTASE.	DR	248 MDYIITDITSAVHALYPLGLKLVNGKLIALGSLPEPELFWFPLIILGRKVAGSDVGMK 307
KW	38934 MW: 0E4F1B0581785759 CRC64;	DR	277 ETQEMIDFAAHNITDADIEVPLDLYNTAMERVKKDQVFRFVLDVENT 326
KW	SEQUENCE 360 AA;	DR	308 ETQEMIDFCAKHNTIADIELKXMDINTAMERLAKSDVTRFVIDVANSI 357
DR	50.1%; Score 81; DB 10; Length 360;	DR	RESULT 8
DR	49.1%; Pred. No. 3. 6e-72; Indels 50; Gaps 5;	DR	055621 PRELIMINARY; PRT; 363 AA.
DR	Matches 172; Conservative 51; Mismatches 77; Indels 50; Gaps 5;	DR	065621 AC 065621: 07. Created
DR	34 DNGENDVTKILFGCVTDLHTKWDGTSYTFPPGHDVIGATKVGNVTRPKGDR 93	DR	01-AUG-1998 (TREMBLrel. 07. Last sequence update)
DR	1 ETGATDVRPKLYCCCHSDHIMAKNDWGRSTYPTVPGHLYGVYTEVGCKYKKPKSWRQ 60	DR	01-OCT-2003 (TREMBLrel. 25. Last annotation update)
DR	34 DNGENDVTKILFGCVTDLHTKWDGTSYTFPPGHDVIGATKVGNVTRPKGDR 93	DR	Cinnamyl alcohol dehydrogenase-like protein, subunit A (Cinnamyl alcohol dehydrogenase-like protein, LCAD), (putative alcohol dehydrogenase) (EC 1.1.1.15).
DR	61 -----GRCWLHGRRLRPTCENCIFHLENYCPLQITYGSKYDSTMVCGSYNNMVYDDE 113	DR	LCADb OR F20510 OR F20510 OR CAD6.
DR	94 VGVGVISGSC-----QSCESCPDQDLBNYCPNSCTYNSGDSDTRKNSYNSVVDQ 146	DR	Arabidopsis thaliana (Mouse ear cress);
DR	114 HETRIPDNFLDGAAPLCLAGITVSPWRYGLDQKPMAMHSPVRRFRSRPPLNPGL 173	DR	Fukuyoya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; rosids; OC
DR	147 RFLVRLPENLPSDSCAAPLCLAGITVSPWRYGLDQKPMAMHSPVRRFRSRPPLNPGL 173	DR	OC euroids II; Brassicales; Brassicaceae; Arabidopsis; OC
DR	174 WGSRLQSLVPLPLKCGSYGTSPALMHS-----LRLRQDQMEAMST 216	DR	RESULT 7
DR	193 GG-----LGHAVVIGKQASLKVYISSSSSTKAEAHINHISADSLTVTIDQKMKAKIGT 247	DR	055620 PRELIMINARY; PRT; 363 AA.
DR	217 MDGIIIDTVPLRPLISLKLTKVWVYQIOPDLFWFPLIILGRKVAGSAIGGMK 276	DR	055620 ID 055620: 22. Created
DR	248 MDYIITDITSAVHALYPLGLKLVNGKLIALGSLPEPELFWFPLIILGRKVAGSDVGMK 307	DR	01-OCT-2002 (TREMBLrel. 22. Last sequence update)
DR	277 ETQEMIDFAAHNITDADIEVPLDLYNTAMERVKKDQVFRFVLDVENT 326	DR	01-OCT-2002 (TREMBLrel. 22. Last sequence update)
DR	308 ETQEMIDFCAKHNTIADIELKXMDINTAMERLAKSDVTRFVIDVANSI 357	DR	01-OCT-2003 (TREMBLrel. 25. Last annotation update)
DR	360 AA;	DR	Cinnamyl alcohol dehydrogenase CAD5.
DR	360 AA;	DR	Arabidopsis thaliana (Mouse ear cress);
DR	360 AA;	DR	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; rosids; OC

OX	NCBI_TaxID=3702;	Db	302 SGIGGNGQETOEMIDPAEHHGKABIEIIISMDYVNTAMDRLLAKGDYRVRVIDISNTL 358
RN	SEQUENCE FROM N.A.		
RC	STRAIN=COLUMBIA, and cv. CO;		
RA	Tavares R., Kreis M.; Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.		
[1]			
RL			
RN			
[2]			
RP	Barvan M., Wedler H., Kurzner M., Wambutt R., Banroft I., Mewes H.W., Mayer K.F.X., Schueler C.; Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.		
RA			
RA			
RA			
RN			
[3]			
RP	SEQUENCE FROM N.A.		
RA	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.		
RN			
[4]			
RP	SEQUENCE FROM N.A.		
RA	Wedler H., Kutzner M., Wambutt R., Mewes H.W., Lemcke K., Mayer K.F.X.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
RN			
[5]			
RP	SEQUENCE FROM N.A.		
RA	EU Arabidopsis sequencing project; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
RN			
[6]			
RP	SEQUENCE FROM N.A.		
RA	Kim S.J., Kim M.R., Bedgar D.L., Moimuddin S.G.A., Cardenas C.I., Davin L.B., Lewis N.G.; "Functional Reclassification of the Putative Cinnamyl Alcohol Dehydrogenase (CAD) Multigene Family in Arabidopsis"; Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.		
RN			
CC	ZINC (BY SIMILARITY).		
DR	EMBL: Y16848; CRA76418.1; -;		
DR	EMBL: AL035538; CAB37537.1; -;		
DR	EMBL: AL16152; CAB042.1; -;		
DR	EMBL: AY302075; AAP59428.1; -;		
DR	PIR: T05624; T05624; -;		
DR	GO: GO:0016491; F:alcohol dehydrogenase activity, zinc-dependent; IEA.		
DR	GO: GO:008270; F:zinc ion binding; IEA.		
DR	InterPro: IPR002055; Adh_zn family.		
DR	PFAM: PF00107; Adh_zinc_N_1.		
DR	GO: GO:004044; F:alcohol dehydrogenase activity, zinc-dependent; IEA.		
DR	GO: GO:008270; F:zinc ion binding; IEA.		
DR	InterPro: IPR002055; Adh_zn family.		
DR	PFAM: PF00107; Adh_zinc_N_1.		
KW	Metal-binding; Oxidoreductase; Zinc.		
SEQUENCE	353 AA; 39017 MW; 3AP82E2D2D9CE3 CRC64;		
SQ			
Query Match	48.5%; Score 853; DB 10; Length 363;		
Best Local Similarity	47.9%; Pred. No. 1.3e-69;		
Matches	171; Conservative 50; Mismatches 72; Indels 64; Gaps 6;		
Qy	1 ETGATDVPEKVLXCGVCHSDTHAKNDGCTSTPIVPGHESLGVYTYEVGCKVKFK---	56	
Db	35 KTGHEEVVTKVLYCGICHSIDLHCKNNEWHSSTIVPLVPGHESLGVYTYEVGCKVKFK---	94	
Qy	57 -----SWRQGRCWHLGRRLPTCENCITHLENVCPNLIQTKGSKYDGMHGGYGRN 108		
Db	95 VGYGCTIDSCR-----TCESEEDQENYCTKATATVGHDDGTTQGYSDH 142		
Qy	109 MTTDEHFTVTRIPDNPLDGAAPLLCAGITTPSPWRYGKDKPGMHLGTYEMPRFRSRPL 168		
Db	143 IVVDERAYKVIPTFLPLVSAAPLLCAGTSMYSPMKYFGTGPDKHVG1-----V 191		
Qy	169 NLPLW-GSRLQSLVPPPLKKEGSSYGPSPALHS-----LIRTDQDQ 209		
Db	192 GLSGLGHGTVFREAK-----AFGTVKTVSSSTGSKDADLTGADGFLVSTDHQ 241		
Qy	210 MEAMSTNDGIDLIDVPAVPLPEPLSLKNGKWTGIVTAQPLDPVPLTIGKMYAG 269		
Db	242 MKLAMGTMDDGIDTVSASHISPLGLXNGKWTGIVTAQPLDPVPLTIGKMYAG 301		
Qy	270 SAIGSMKTOEMIDFAAEHNITADIEVPIPDYINTAMERVKDKVRFVFDVENTL 326		
		RESULT 9	
		SEQUENCE FROM N.A.	
Q8S411	AC Q8S411; PRELIMINARY; PRT; 370 AA.		
ID	DT 01-JUN-2002 (TREMBLrel. 21, Created)		
AC	DT 01-JUN-2003 (TREMBLrel. 21, Last annotation update)		
Q8S411	DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
	DE Cinnamyl alcohol dehydrogenase.		
	GN CAD2.		
	OS Lolium perenne (perennial ryegrass).		
	OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Poae; Lolium.		
	NCBI_TaxID=4522;		
	OX RN [1] _taxID=4522;		
	RP SEQUENCE FROM N.A.		
	RC STRAIN=CV; Elet;		
	RA Lynch D., Liggett A., McLaren R., Huxley H., Jones E., Mahoney N.,		
	RA Spangenberg G.;		
	RT "Isolation and Characterization of Three Cinnamyl Alcohol Dehydrogenase Homolog cDNAs from Perennial Ryegrass [loliu...L.]".		
	RT J. Plant Physiol. 0:0-0 (2002).		
	RL CC J. COFACTOR; ZINC (BY SIMILARITY).		
	DR EMBL; AF472592; AAU9536.1; -;		
	DR GO: GO:000424; F:alcohol dehydrogenase activity, zinc-dependent; IEA.		
	DR GO: GO:0005489; F:electron transporter activity; IEA.		
	DR GO: GO:0016490; F:oxidoreductase activity; IEA.		
	DR GO: GO:0008270; F:zinc ion binding; IEA.		
	DR GO: GO:0006118; F:electron transport; IEA.		
	DR InterPro; IPR02328; Adh_zn.		
	DR InterPro; IPR02085; Adh_zn family.		
	DR InterPro; IPR00345; CytC_heme_BS.		
	DR Pfam; PF00107; Adh_zinc_N_1.		
	DR PROSITE; PS00359; ADH_ZINC_1.		
	DR PROSITE; PS00190; CYTOCHROME_C_1.		
	KW Metal-binding; Oxidoreductase; Zinc.		
SQ	SEQUENCE 370 AA; 39390 MW; 9058D20FA9096E47 CRC64;		
	Query Match 47.7%; Score 839; DB 10; Length 370;		
	Best Local Similarity 49.0%; Pred. No. 2.6e-68;		
	Matches 169; Conservative 51; Mismatches 86; Indels 38; Gaps 5;		
Qy	2 TGTATVRFKVLXCGVCHSDTHAKNDGCTSTPIVPGHESLGVYTYEVGCKVKFKW-RQ 60		
Db	40 TGDDWVVKILYCGICHSIDLHCKNNEWHSSTIVPLVPGHESLGVYTYEVGCKVKFKW-RQ 99		
Qy	61 GRCWTHGRKPTCNCNTHLRYCNPNLIQTYGSKYDGMHGGYGRN 120		
Db	100 GVGGMVNSCR-SCSCKDGFENHCPGMILTYNSVSDVDTGTVTYGSKYDGMHGGYGRN 158		
Qy	121 DNLDGAPLICGIGTTSPWRYGLDKPGMHLGTYEWPRFRSRPLPGLGMSRLQS 180		
Db	159 DAMPDKGAPLICGIGTTSPWRYGLDKPGMHLGTYEWPRFRSRPLPGLGMSRLQS 201		
Qy	181 LVPPPLIKEGSSYGPSPALHS-----LIRTDQDQ 223		
Db	202 -HYAVKFGKAFGMKTVTSSSPKKEEALGRLGADAFIVSKDADMVKAVMSMDINT 259		
Qy	224 VPAYPLPLWLIKNGKWTGIVTAQPLDPVPLTIGKMYAG 283		
Db	260 VSANPLPLFGLKPGHGMVOLPKEPEIPFALVNTKTLAGSITGGMSDTPQEMD 319		
Qy	284 FAABHNITADIEVPIPDYINTAMERVKDKVRFVFDVENTL 326		
Db	320 LAAGHGVADIEVGAETVNTAERLAKNDVYTFVFDVENTL 362		
	RESULT 10		
	Q9S510		

1	Q9SJU0	PRELIMINARY;	PRT;	375 AA.	
AC	Q9SJU0; (1)				
DT	01-MAY-2000 (TREMBL); 13; Created)				
DT	01-MAY-2000 (TREMBL); 13; Last sequence update)				
DT	01-OCT-2003 (TREMBL); 25; Last annotation update)				
DN	Cinnamyl alcohol dehydrogenase-like protein (EC 1.1.1.195).				
GN	ATG21890 OR CAD3.				
OS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
OX	NCBI_TaxID=3702;				
RN					
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Cv. Columbia;				
RA	MEDLINE=20083487; PubMed=10617197;				
RA	Lin X., Kau S., Ronning S.D., Shea T.P., Benito M.-I., Town C.D., Fujii C.Y., Mason T.M., Bowman C.I., Barnstead M.E., Fieldblyum T.V., Bueli C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S., Cronin L.A., Vanek S.E., Umayam L., Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R., Copenhafer G.P., Preuss D., Nierman W.C., White C., Eisen J.A., Salzberg S.L., Fraser C.M., Venter J.C.;				
RA	"Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana"; Nature 402:761-768 (1999).				
RA	[2]				
RA	SEQUENCE FROM N.A.				
RA	STRAIN=Cv. Columbia;				
RA	Lin X.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.				
RA	Kim S.J., Kim M.R., Bedgar D.L., Moinuddin S.G.A., Cardenas C.L., Davin L.B., Lewis N.G.; "Functional Reclassification of the Putative Cinnamyl Alcohol Dehydrogenase (CAD) Multigene Family in Arabidopsis"; Submitted (MAY-003) to the EMBL/GenBank/DBJ databases.				
RA	-1- COFACTOR: ZINC (BY SIMILARITY).				
RA	EMBL; AX302078; AAP9431.1; -.				
RA	EMBL; AX302078; AAP9431.1; -.				
RA	PIR; D84606; D84606.				
DR	GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.				
DR	GO:00016491; F:oxidoreductase activity; IEA.				
DR	GO:00038270; F:zinc ion binding; IEA.				
DR	InterPro; IPR002328; ADH_zinc.				
DR	InterPro; IPR002085; Adh_zn_family.				
DR	PFam; PF0107; ADH_zinc_N; 1.				
DR	Metal-binding; Oxidoreductase; Zinc.				
DR	SEQUENCE: 375 AA; 40886 MW; ADC1FFCABEA2D113 CRC64;				
DR					
DR	Query Match	46.6%	Score 819.5;	DB 10;	Length 375;
DR	Best Local Similarity	47.6%	Pred. No. 1.6e-16;		
DR	Matches	167; Conservative	48; Mismatches	85; Indels	51; Gaps 6;
DR					
DR	82	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	113		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIGT-----QCSQCNOLENPKVPLKIGT-----	186		
DR	187	LGG-----LGHIAVKGKGLRVTVISRSSEKERAIDRIGADSFLVTISQRMKEAVG	241		
DR	28	ENGENDVTKILPCGVCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	29	1 ETGATDVRKFLYCYCHSDLHTKKNHWGFSRYP1PGHE1ATKVGNVTKFEGDR	87		
DR	88	GGVGTIGS-----QCSQCNOLENPKVPLKIGT-----	140		
DR	114	HEIVRIPDNLPDGAAPLCAPILGCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	141	RFVLSIPTDGPSPDSGAPLCACTTSPWRYGLDK-PGMILGVENPRFRSPRPLNLPG	172		
DR	173	LWGSRLQLSVPPLKIG			

Qy	113	EHPFVRIEPDNLPDGAAPLLCAGITYSPPWRYGLDK-PGMHLGVENPRFRSRPPLNP	171	Db	372	LAAXHGGTADIEVYGGADDYNTAMERLAKADVYRFVIDVGNTL	414
Db	141	HRFVLSIEDGLISDSGAPLLCAGITYSPPWRYGMTEKSGRIGVN-----	186				
Qy	172	GHWSRILQSLVPPLEKRGSGYGTSPALMHS-----	214	RESULT 13			
Db	187	GLGG-----LGHIAVAKIGRAFLGLRVTVTSSSEKERAIDRGLADSPFLVTTDSQKMKRAV	241	QSM722	PRELIMINARY;		
Db	215	STMDGIDTTPVAPVRPLPLISLJKTNGKVTGIAVQPLDLPVPLIGRNGVAGSIGG	274	ID QSM722;			
Qy	242	GTMDFIDTDSVSEAHLLPLFLSFLKVKXLLVALGFLPFLFISLVLYGRNGGSGQIG	301	DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
Db	275	MKETORMIDFAEHNITADIEVTPIDVNTAMERVVKKDVPRFVIDVENTL	326	DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
Qy	302	MKETORMIDFAEHNITADIEVTPIDVNTAMERVVKKDVPRFVIDVENTL	326	DE ELI3 (Fragment).			
Db	302	MKETORMIDFAEHNITADIEVTPIDVNTAMERVVKKDVPRFVIDVENTL	326	OS Lycopersicon esculentum (Tomato).			
				OC Bokarysta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
				OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;			
				OC Lamiales; Solanales; Solanaceae; Solanum.			
				OX NCBI_TaxID=4081;			
				RN [1]			
				RP SEQUENCE FROM N.A.			
				RC STRAIN=cv. Rio Grande;			
				RA Thiburney, P. L.; Martin, G. B.;			
				RT "Rapid Induction of a Novel Gene and an eli3 Homolog During the Pto- mediated Resistance Response in Tomato."			
				RT Submitted (APR-1999) to the EMBL/GenBank/DDJB databases.			
				DR EMBL; AF146591; AAF72100.1; -.			
				DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.			
				DR GO; GO-0008270; F:zinc ion binding; IEA.			
				DR InterPro; IPR00085; Adh_zn_family.			
				DR Pfam; PF00107; ADH_zinc_N; 1.			
				FT NON_TER 1			
				SQ SEQUENCE 289 AA; 31333 MW; B2BAPDFC316A66CE8 CRC64;			
				Query Match 44.6%; Score 784.5; DB 10; Length 299;			
				Best Local Similarity 52.4%; Pred. No. 1..8e-33;			
				Matches 161; Conservative 50; Mismatches 57; Indels 39; Gaps 6;			
Qy	3	GHELVGVVTEVGCKVFKF-SWRCGCKWLRGLRPPCENCYTHLETCPNLJOTYGSKYY	96	38 GHELVGVVTEVGCKVFKF-SWRCGCKWLRGLRPPCENCYTHLETCPNLJOTYGSKYY	96		
Db	1	GTRVRYGVTEVGSKVKPKVCKVFKF-SWRCGCKWLRGLRPPCENCYTHLETCPNLJOTYGSKYY	96	97 DGTMTTGGYSNNMVTDFHFTVTRIPONPLDGAAPLICAGTTYSPPRYYGLDKPGHNLGV	156		
Db	60	DGTLLTGGTSSVMSSBHFVFRWPELMSD-AAPLICAGTTYSPPRYYGLDKPGHNLGV	118	60 DGTLLTGGTSSVMSSBHFVFRWPELMSD-AAPLICAGTTYSPPRYYGLDKPGHNLGV	118		
Qy	157	EWPRFRSRPPLNLPCLWGSPLQSLVPLIKEGGSGTSPALMS-----	231	157 EWPRFRSRPPLNLPCLWGSPLQSLVPLIKEGGSGTSPALMS-----	231		
Db	119	-----VGLGG-----LGCHMAYTKFAKAFGTVKTVLTSRANKKKEAISRLGAD	159	Db 119 -----VGLGG-----LGCHMAYTKFAKAFGTVKTVLTSRANKKKEAISRLGAD	159		
Qy	202	LLRUDQDMELAATSTMDGILDTTPAVRPLPLSLLKINGKVVYGVIAVQLPDPVFP	259	Qy 202 LLRUDQDMELAATSTMDGILDTTPAVRPLPLSLLKINGKVVYGVIAVQLPDPVFP	259		
Db	160	SPLISRDPRKMKATNLGIDTYVSSVHLPLPLMMKSGKLVPMGAPEKPVFVFP	219	Db 160 SPLISRDPRKMKATNLGIDTYVSSVHLPLPLMMKSGKLVPMGAPEKPVFVFP	219		
Qy	260	LIGGRMVAAGMKEDEMIDPAEINTTADIEVTPIDYLTANMERYVKKDGYFTRFV	319	Qy 260 LIGGRMVAAGMKEDEMIDPAEINTTADIEVTPIDYLTANMERYVKKDGYFTRFV	319		
Db	220	LLMGRKLVAGSGCIGMKETOEMLDFAKAHTPDEVMEMEVNTAELRJSDVYRFV	279	Db 220 LLMGRKLVAGSGCIGMKETOEMLDFAKAHTPDEVMEMEVNTAELRJSDVYRFV	279		
				Qy 320 IDVENTL 326			
				Db 280 LDGNTL 286			
				RESULT 14			
Qy	181	LVPPLIKEGGSGTSPALMHS-----LIRTIDQDMBAMSTMDGIDT	223	Q7XW00 PRELIMINARY;			
Db	254	-HVAKFARAFELKVTUSSSPGKKGREALERLGADEFVSSAEMRSTMDGVNT	311	ID Q7XW00;			
Qy	224	VPAVRPLPLISLJKTNGKVTGIAVQPLDLPVPLIGRNGVAGSAGMKTQEMD	283	DT 01-OCT-2003 (TREMBLrel. 25, Created)			
Db	312	VSANTPMAYLALLKPNGLNIVLGPFLPENLVEPPSLVNGNRTLAGSNIGMADTQEME	371	DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)			
Qy	284	FAAEHNITADIEVTPIDYNTAMERVVKKDVPRFVIDVENTL	326	DE OSJNBA0065B15.11 protein.			
				GN OSJNBA0065B15.11.			
				OS Oryza sativa (Rice).			
				OC Bokarysta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
				OC Liliopsida; Poales; Poaceae;			
				OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			

OC	Ehrhartoideae; Oryzeae; Oryzae.	DR	InterPro; IPR002085; Adh_zinc family.
OX	NCBI_TaxID=4530;	DR	InterPro; IPR00035; Cyt_C_heme_BS.
RN	SEQUENCE FROM N.A.	DR	Pfam; PF00107; Adh_zinc_N_1.
RP		DR	PROSITE; PS00059; Adh_zinc_N_1.
RA	Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,	DR	PROSITE; PS00190; CYTOCHROME_C_1.
RA	Lu Y.Q., Mu L., Chen J., Chen L., Fan D.L., Zhang D.,	SQ	SEQUENCE 391 AA; 40917 MW; EB3D9978B543C55B CRC64;
RA	Ren S.X., Lv G., Lin W., Liu X.H., Liu T.T., Zhang Y.J., Lu Y.,	Query Match	40.5%; Score 712.5; DB 10; Length 391;
RA	Gu G., Wang S.Y., Guan J.P., Fu G., Wang S.Y.,	Best Local Similarity	40.6%; Pred. No. 1.1e-56;
RA	Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,	Matches 154; Conservative 55; Indels 93; Gaps 7;	Matches 154; Conservative 55; Indels 93; Gaps 7;
RA	Ren Q.P., Zhang L., Wu N., Zhang R.Q., Gu W.Q., Zhou G.F., Tu Y.F., Jia J.,	QY	3 GATIVPRKVLYCCYCHSDITHMAKNDWTFSTYPIVPGHLLVGIVTEVGCKVKKPKSWRQCR 62
RA	Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,	Db	31 GDEDWVKTILFCGICHSQDISTINNEAKYKPVPGHIVGVAEVGSSVARPA--GD 87
RA	Zhang X.L., Ni L., Zhu W., Wang L.J., Ding C.W., Sheng H.H.,	QY	63 CWLHGRLRPTCE--NCLHHLNCPNLQITQGSKYIDGTMITYGGYNNMVTDEHFTYRI 119
RA	Gu J.L., Chen S.T., Hong G.F.,	Db	88 TVGVGYIATSTCRACANCGRDFENYCAGLVPSENVALPGATVHGGFSELAVINQRYVTRI 147
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.	Db	120 PDN----LPIDGAAPILCAPIGTTSPWRYTGLDKPGMELAVEMPFRSRPLNLPG 173
DR	EMBL; AL731598; CAD39307.1; -;	QY	148 PGGGGASPAIDLRLAPLICGTVYCMRRLGDRPQHLYT- AGL 193
DR	NCBI; 39086 MW; ODEA56D0DBB425CA CRC64;	Db	174 WGSRLQSLVPPPLIKEGGSYGTSPALMS-----LIRTDDQOMEAMST 216
SEQUENCE	360 AA;	QY	Db 194 GG----LGHLLAVKRGAFGKAVTIVTSPWKEAEAVRLGDAFLSTNAEOKKAAGT 248
SEQUENCE	360 AA;	Db	217 MDGIDTDPAVRPLPEPLISLKLTKNGKWTGAVQPLDPVPLP----- 260
Best Local Similarity	43.0%; Score 755.5; DB 10; Length 360;	QY	Db 249 MDGIDTDPASVADLTLPLTLLRTHGQLEWVGSPKGPKQLALYPLQSPSPVGRIDWGG 308
Matches 160; Conservative 45; Mismatches 96; Indels 43; Gaps 8;	Db	261 -----IIGKNAVASSAIGSKETOBMIDAAEINNTADDEVIPIDYINTAME 307	
Qy	1 ETGATDVRFKVLPGVCHSDITHMAKNDWTFSTYPIVPGHLLVGIVTEVGCKVKKPK-SWR 59	Db	309 PIRARWAVLMLCDVADGSKVAGSMIGMARDQEMDPAVEVIGHEDRSGAME 368
Db	32 ENGDDVTVIKVYCGMHTDLEIFIDNDGTMVPPVFGHEITGVVKGTVNAGFKVGDR 91	QY	308 RVVKKDYRERFVLDVENTL 326
Qy	60 QG-RCWPLHGRLRPTCNCIHLHENYCPNLQITQGSKYDGTMTYGGYSNNAYTDEHFTIVR 118	Db	369 RLQGDYRTRFYDVENTM 387
Db	92 VGVGCTAASCL--DCEFCRRSEBNYCDKVALITYNGTFWDGTSITYGGYSCMVAHKRFVVR 149	QY	Search completed: July 6, 2004, 13:38:47
Qy	119 LPDNLPLDGAAPPLCAGITTYSPWRYYGL-- DKGPMHLLGIEWPFRFRSPRPLNLPLGLWG 175	Db	Job time : 33.8458 secs
Db	150 LPDTPLDAAAPPLCAGITVSPMKQHGMQADAASRRLGV-----VGLGG 195	QY	
Qy	176 SRLQSLVPPPLIKEGGSYG-----TSPALMS-----LIRTDDQOMEAMSTD 218	Db	
Db	196 -----LGHVATKFGKAFGLHVTVISTSPAKEREARENLKADNEVISTDQKMQAMTRSLD 250	QY	
Qy	219 GLIDTYPAVPLLEPLISLKLNGKWTGAVQPLDPVPLIIGKRMVAGSAIGMKET 278	Db	
Db	251 YIIDTVAAHTSPLGPBELLKGKLVLGAEKPVLPSPFLLGFRTVSSMTGKET 310	QY	
Qy	279 QEMIDFAAEHNITADIEVTPIDLYNTAMERYVKKVDRFVTFIDV 322	Db	
Db	311 QEMMDIGEHNITCDIBVSTDRINDALARLARDVYRFTINV 354	RESULT 15	
OS	Oryza sativa (Japonica cultivar-group).	QBBH809	PRELIMINARY; PRT; 391 AA.
ID		QBBH809	
AC		QBBH809;	
DT	01-MAR-2003 (TREMBLrel. 23, Created)	QBBH809;	
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)	QBBH809;	
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	QBBH809;	
DE	Putative mannitol dehydrogenase (NAD-dependent mannitol dehydrogenase).	QBBH809;	
DE	QBBH809.18.	QBBH809;	
OS	Oryza sativa (Japonica cultivar-group).	QBBH809	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	QBBH809	
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	QBBH809	
OC	Ehrhartoideae; Oryzeae; Oryzae.	QBBH809	
QX	NCBI_TaxID=39947;	QBBH809	
RN	SEQUENCE FROM N.A.	QBBH809	
RA	Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Sasaki C.,	QBBH809	
RA	Currie J., Collura K.;	QBBH809	
RT	"Rice Genomic Sequence."	QBBH809	
DR	Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.	QBBH809	
DR	EMBL; AC105364; AAC05330.1;	QBBH809	
DR	GO; GO:0004024; F:alcohol dehydrogenase activity; IEA.	QBBH809	
DR	GO; GO:0005489; F:electron transporter activity; IEA.	QBBH809	
DR	GO; GO:0008270; F:finc ion binding; IEA.	QBBH809	
DR	GO; GO:0006118; P:electron transport; IEA.	QBBH809	
DR	InterPro; IPR002328; Adh_zinc.	QBBH809	



Creation date: 07-14-2004

Indexing Officer: MMOHAMMED1 - MENEN MOHAMMED

Team: OIPEScanning

Dossier: 09857518

Legal Date: 07-09-2004

No.	Doccode	Number of pages
1	SRNT	26

Total number of pages: 26

Remarks:

Order of re-scan issued on